



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : A01H 1/00, C07H 21/04, C07K 14/00, C12N 5/04, 5/10, C12P 19/34, C12Q 1/68	A1	(11) International Publication Number: WO 98/30083 (43) International Publication Date: 16 July 1998 (16.07.98)
(21) International Application Number: PCT/US98/00615 (22) International Filing Date: 9 January 1998 (09.01.98) (30) Priority Data: 08/781,734 10 January 1997 (10.01.97) US (71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612 (US). (72) Inventors: SHEN, Kathy; 44228 Country Club Drive, El Macero, CA 95618 (US). MEYERS, Blake; 904 Drake Drive, Davis, CA 95616 (US). MICHELMORE, Richard, W.; 36757 Russel Boulevard, Davis, CA 95616 (US). (74) Agents: EINHORN, Gregory, P. et al.; Townsend and Townsend and Crew LLP, 8th floor, Two Embarcadero Center, San Francisco, CA 94111 (US).		(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: RG NUCLEIC ACIDS FOR CONFERRING DISEASE RESISTANCE TO PLANTS (57) Abstract The present invention provides RG nucleic acids and proteins which confer disease resistance to plants. The nucleic acids can be used to produce transgenic plants resistant to pests. Antibodies to proteins of the invention are also provided.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

RG NUCLEIC ACIDS FOR CONFERRING DISEASE RESISTANCE TO PLANTS

5

The present application is a continuation-in-part application ("CIP") of U.S. Patent Application Serial No. ("USSN") 08/781,734, filed January 10, 1997. The
10 aforementioned application is explicitly incorporated herein by reference in its entirety and
for all purposes.

This invention was made with Government support under Grant Nos. 92-37300-7547 and 95-37300-1571, awarded by the United States Department of Agriculture.
15 The Government has certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates generally to plant molecular biology. In particular, it relates to nucleic acids and methods for conferring pest resistance in plants.
20 particularly lettuce.

BACKGROUND OF THE INVENTION

Recently, several resistance genes have been cloned by several groups from several plants. Many of these genes are sequence related. The derived amino acid
25 sequences of the most common class, *RPS2*, *RPM1* (bacterial resistances in *Arabidopsis* (Mindrinos *et al.* *Cell* 78:1089-1099 (1994)); Bent *et al.* *Science* 265:1856-1860 (1994); Grant *et al.*, *Science* 269:843-846 (1995)), *L6* (fungal resistance in flax; Lawrence, *et al.*, *The Plant Cell* 7:1195-1206 (1995)), and *N*, (virus resistance in tobacco; Whitham, *et al.*, *Cell* 78:1101-1115 (1994); and U.S. Patent No. 5,571,706), all contain leucine-rich
30 repeats (LRR) and nucleotide binding sites (NBS).

The NBS is a common motif in several mammalian gene families encoding signal transduction components (e.g., *Ras*) and is associated with ATP/GTP-binding sites.

The NBS is a common motif in several mammalian gene families encoding signal transduction components (e.g., *Ras*) and is associated with ATP/GTP-binding sites.

LRR domains can mediate protein-protein interactions and are found in a variety of proteins involved in signal transduction, cell adhesion and various other functions. LRRs are leucine rich regions often comprising 20-30 amino acid repeats where leucine and other aliphatic residues occur periodically. LRRs can function extracellularly or intracellularly.

Since the onset of civilization, plant diseases have had catastrophic effects on crops and the well-being of the human population. Plant diseases continue to effect enormous human and economic costs. An increasing human population and decreasing amounts of arable land make all approaches to preventing and treating plant pathogen destruction critical. The ability to control and enhance a plant's protective responses against pathogens would be of enormous benefit. Tissue-specific and temporal control of mechanisms responsible for plant cell death would also be of great practical and economic value. The present invention fulfills these and other needs.

What is needed in the art are plant disease resistance genes and means to create transgenic disease resistance plants, particularly in lettuce. Further, what is needed in the art is a means to DNA fingerprint cultivars and germplasm with respect to their disease resistance haplotypes for use in plant breeding programs. The present invention provides these and other advantages.

SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acid constructs. These constructs comprise an RG (resistance gene) polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, and an RG4 polypeptide. RG1, RG2, RG3, RG4, and the like, represent individual "RG families." Each "RG family," as defined herein, is a group of polypeptide sequences that have at least 60% amino acid sequence identity. Individual members of an RG family, i.e., individual species of the genus, typically map to the same genomic locus. The invention provides for constructs comprising nucleotides encoding the RG families of the

invention, which can include sequences encoding a leucine rich region (LRR), and/or a nucleotide binding site (NBS), or both.

The invention provides for an isolated nucleic acid construct comprising an RG polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide from an RG family selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide. In alternative embodiments, the nucleic acid construct comprises an RG polynucleotide which encodes an RG polypeptide comprising an leucine rich region (LRR), or, an RG polypeptide comprising a nucleotide binding site (NBS). The nucleic acid construct can comprise a polynucleotide which is a full length gene. In another embodiment, the nucleic acid construct encodes a fusion protein.

In one embodiment, the nucleic acid construct comprises a sequence encoding an RG1 polypeptide. The RG1 polypeptide can be encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 and SEQ ID NO:137 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).

In another embodiment, the nucleic acid construct comprises a sequence encoding an RG2 polypeptide. The RG2 polypeptide can be encoded by a polynucleotide sequence selected from the group consisting of: SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

In other embodiments, the nucleic acid construct comprises a RG3 sequence (SEQ ID NO:68) encoding an RG3 polypeptide (SEQ ID NO:138) (RG3). In other embodiments, the nucleic acid construct comprises an RG4 sequence (SEQ ID NO:69) encoding an RG4 polypeptide (SEQ ID NO:139) (RG4).

5 In other embodiments, the nucleic acid construct comprises a RG5 sequence (SEQ ID NO:134) encoding an RG5 polypeptide (SEQ ID NO:135). The RG5 polypeptide can be encoded by a polynucleotide sequence as set forth in SEQ ID NO:134.

The invention also provides for a nucleic acid construct which comprises an RG7 sequence encoding an RG7 polypeptide. The RG7 polypeptide can be encoded by a polynucleotide sequence as set forth in SEQ ID NO:136.

10 In further embodiments, the nucleic acid construct can further comprise a promoter operably linked to the RG polynucleotide. In alternative embodiments, the promoter can be a plant promoter; a disease resistance promoter; a lettuce promoter; a constitutive promoter; an inducible promoter; or, a tissue-specific promoter. The nucleic acid construct can comprise a promoter sequence from an RG gene linked to a heterologous polynucleotide.

The invention also provides for a transgenic plant comprising a recombinant expression cassette comprising a promoter operably linked to an RG polynucleotide. The expression cassette can comprise a plant promoter or a viral promoter; the plant promoter can be a heterologous promoter. In one embodiment, the transgenic plant is lettuce. In alternative embodiments, the transgenic plant comprises an expression cassette which includes an RG polynucleotide selected from the group consisting of SEQ ID NO:1 (RG1A); SEQ ID NO:2 and SEQ ID NO:137 (RG1B); SEQ ID NO:3 (RG1C); SEQ ID NO:4 (RG1D); SEQ ID NO:5 (RG1E); SEQ ID NO:6 (RG1F); SEQ ID NO:7 (RG1G); SEQ ID NO:8 (RG1H); SEQ ID NO:9 (RG1I) and SEQ ID NO:10 (RG1J); SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104

(RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W); SEQ ID NO:68 (RG3); SEQ ID NO:69 (RG4); SEQ ID NO:134 (RG5); or SEQ ID NO:136 (RG7).

The invention provide for a transgenic plant comprising an expression cassette comprising an RG polynucleotide which can encode an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO:13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), or SEQ ID NO:20 (RG1J); or, an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W); an RG4 polypeptide as set forth by SEQ ID NO:72; an RG5 polypeptide with a sequence as set forth by SEQ ID NO:135; or, an RG7 polypeptide.

The invention also provides for a method of enhancing disease resistance in a plant, the method comprising introducing into the plant a recombinant expression cassette comprising a promoter functional in the plant and operably linked to an RG polynucleotide sequence. In this method, the plant can be a lettuce plant; and, the RG polynucleotide can encode an RG polypeptide selected from the group consisting of an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO:13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID

NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), or SEQ ID NO:20 (RG1J); or, an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:72; SEQ ID NO:74; SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W). In this method, the promoter can be a plant disease resistance promoter, a tissue-specific promoter, a constitutive promoter, or an inducible promoter.

The invention also provides for a method of detecting RG resistance genes in a nucleic acid sample, the method comprising: contacting the nucleic acid sample with an RG polynucleotide to form a hybridization complex; and, wherein the formation of the hybridization complex is used to detect the RG resistance gene in the nucleic acid sample. In this method, the RG polynucleotide can be an RG1 polynucleotide, an RG2 polynucleotide, an RG3 polynucleotide, an RG4 polynucleotide, an RG5 polynucleotide or an RG7 polynucleotide. In this method, the RG resistance gene can be amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide, and, the RG resistance gene can be amplified by the polymerase chain reaction. In one embodiment, the RG polynucleotide is labeled.

The invention further provides for an RG polypeptide having at least 60% sequence identity to a polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification, the figures and claims.

All publications, patents and patent applications cited herein are hereby expressly incorporated by reference for all purposes.

DETAILED DESCRIPTION OF THE INVENTION

This invention relates to families of RG genes, particularly from *Lactuca sativa*. Nucleic acid sequences of the present invention can be used to confer resistance in plants to a variety of pests including viruses, fungi, nematodes, insects, and bacteria. Sequences from within the RG genes can be used to fingerprint cultivars or germplasm for the presence of desired resistance genes. Promoters of RG genes can be used to drive heterologous gene expression under conditions in which RG genes are expressed. Further, the present invention provides RG proteins and antibodies specifically reactive to RG proteins. Antibodies to RG proteins can be used to detect the type and amount of RG protein expressed in a plant sample.

The present invention has use over a broad range of types of plants, including species from the genera *Cucurbita*, *Rosa*, *Vitis*, *Juglans*, *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersicon*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Ciahorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Heterocallis*, *Nemesis*, *Pelargonium*, *Panieum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Pisum*, *Phaseolus*, *Lolium*, *Oryza*, *Zea*, *Avena*, *Hordeum*, *Secale*, *Triticum*, and *Sorghum*. In particularly preferred embodiments, species from the family *Compositae* and in particular the genus *Lactuca* are employed such as *L. sativa* and such subspecies as *crispa*, *longifolia*, and *asparagina*.

The nucleic acids of the present invention can be used in marker-aided selection. Marker-aided selection does not require the complete sequence of the gene or precise knowledge of which sequence confers which specificity. Instead, partial sequences can be used as hybridization probes or as the basis for oligonucleotide primers to amplify nucleic acid, e.g., by PCR. Partial sequences can be used in other methods, such as to

follow the segregation of chromosome segments containing resistance genes in plants. Because the RG marker is the gene itself, there can be negligible recombination between the marker and the resistance phenotype. Thus, RG polynucleotides of the present invention provide an optimal means to DNA fingerprint cultivars and wild germplasm with respect to their disease resistance haplotypes. This can be used to indicate which germplasm accessions and cultivars carry the same resistance genes. At present, selection of plants (e.g., lettuce) for resistance to some diseases is slow and difficult. But linked markers allow indirect selection for such resistance genes. Moreover, RG markers also allow resistance genes to be identified and combined in a manner that would not otherwise be possible. Numerous accessions have been identified that provide resistance to all isolates of downy mildew (*Bremia lactucae*). However, without molecular markers it is impossible to combine such resistances from different sources. The nucleic acid sequences of the invention provide for a fast and convenient means to identify and combine resistances from different sources. The RG markers of the invention can also be used to identify recombinants that have new combinations of resistance genes in *cis* on the same chromosome.

In addition, RG markers may allow the identification of the Mendelian factors determining traits, such as field resistance to downy mildew. Once such markers have been identified, they will greatly increase the ease with which field resistance can be transferred between lines and combined with other resistances.

In another application, primers to RG sequences can be also designed to amplify sequences that are conserved in multiple RG family members. This gives genetic information on multiple RG family members. Alternatively, one or more primers can be made to sequences unique to a single resistance gene genus or a single RG specie. This allows an analysis of individual family groups (an RG genus) or an individual family member (a specie). Primers made to individual RGs at the edge of each cluster can be used to select for recombinants within the cluster. This minimizes the amount of linkage drag during introgression. Classical and molecular genetics has shown that pest resistance genes tend to be clustered in the genome. Pest resistance loci comprise arrays of genes and exhibit a variety of complex haplotypes rather than being simple alternate allelic forms. Pest resistance is conferred by families, or genuses, of related RG sequences, individual members, or species, of which have evolved to have a different specificity.

Oligonucleotide primers can be designed that amplify members from multiple haplotypes, or genuses, or amplify only members of one genus, or only amplify an individual specie. This will provide codominant information and allow heterozygotes to be distinguished from homozygotes.

5 Further, comparison of RG sequences will allow a determination of which sequences are critical for resistance and will ultimately lead to engineering resistance genes with new specificities. Resistance gene sequences were not previously available for lettuce. Marker-aided selection will greatly increase the precision and speed of breeding for disease resistance. Transgenic approaches will allow pyramiding of resistance genes
10 into a single Mendelian unit, transfer between sexually-incompatible species, substitute for conventional backcrossing procedures, and allow expression of other genes in parallel with resistance genes.

The RG polynucleotides also have utility in the construction of disease resistant transgenic plants. This avoids lengthy and sometimes difficult backcrossing
15 programs currently necessary for introgression of resistance. It is also possible to transfer resistance polynucleotides between sexually-incompatible species, thereby greatly increasing the germplasm pool that can be used as a source of resistance genes. Cloning of multiple RG sequences in a single cassette will allow pyramiding of genes for resistance against multiple isolates of a single pathogen such as downy mildew or against multiple
20 pathogens. Once introduced, such a cassette can be manipulated by classical breeding methods as a single Mendelian unit.

Transgenic plants of the present invention can also be constructed using an RG promoter. The promoter sequences from RG sequences of the invention can be used with RG genes or heterologous genes. Thus, RG promoters can be used to express a
25 variety of genes in the same temporal and spatial patterns and at similar levels to resistance genes.

Nucleic acids of the Invention and Their Preparation

RG Polynucleotide Families

30 The present invention provides isolated nucleic acid constructs which comprise an RG polynucleotide. In alternative embodiments, the RG polynucleotide is at least 18 nucleotides in length, typically at least 20, 25, or 30 nucleotides in length, more

typically at least 100 nucleotides in length, generally at least 200 nucleotides in length, preferably at least 300 nucleotides in length, more preferably at least 400 nucleotides in length, and most preferably at least 500 nucleotides in length.

In particularly preferred embodiments, the RG polynucleotide encodes a RG protein which confers resistance to plant pests. This RG protein can be longer, equivalent, or shorter than the RG protein encoded by an RG gene. In various embodiments, an RG polynucleotide can hybridize under stringent conditions to members of an RG family (an RG genus); *e.g.*, it can hybridize to a member of the RG1 RG family, such as an RG1 polynucleotide selected from the group consisting of: SEQ ID NO:1 (RG1A); SEQ ID NO:2 and SEQ ID NO:137 (RG1B); SEQ ID NO: 3 (RG1C); SEQ ID NO:4 (RG1D); SEQ ID NO:5 (RG1E); SEQ ID NO:6 (RG1F); SEQ ID NO:7 (RG1G); SEQ ID NO:8 (RG1H); SEQ ID NO:9 (RG1I) and SEQ ID NO:10 (RG1J).

In other embodiments, the polynucleotide can also hybridize under stringent conditions to a member of the RG2 family; such as an RG2 polynucleotide selected from the group consisting of: SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

In alternative embodiments, each RG2 gene can also include an AC15 sequence which hybridizes under stringent conditions to a polynucleotide selected from the group consisting of: SEQ ID NO:56 (AC15-2A); SEQ ID NO:57 (AC15-2B); SEQ ID NO:58 (AC15-2C); SEQ ID NO:59 (AC15-2D); SEQ ID NO:60 (AC15-2E); SEQ ID NO:61 (AC15-2G); SEQ ID NO:62 (AC15-2H); SEQ ID NO:63 (AC15-2I); SEQ ID

NO:64 (AC15-2J); SEQ ID NO:65 (AC15-2L); SEQ ID NO:66 (AC15-2N); SEQ ID NO:67 (AC15-2O).

In other embodiments, an RG polynucleotide can hybridize under stringent conditions to an RG3 (SEQ ID NO:68), an RG4 (SEQ ID NO:69), and RG5 (SEQ ID NO:135), and an RG7 (SEQ ID NO:137), RG family member.

The present invention further provides nucleic acid constructs which comprise an RG polynucleotide which encodes RG polypeptides from various RG families; such as an RG polypeptide having at least 60% sequence identity to an RG polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, and RG4 polypeptide, and RG5 polypeptide, and an RG7 polypeptide.

Exemplary RG1 polypeptides have the sequences shown in SEQ ID NO:2 (RG1A), SEQ ID NO:4 (RG1B), SEQ ID NO:6 (RG1C), SEQ ID NO:8 (RG1D), SEQ ID NO:10 (RG1E), SEQ ID NO:12 (RG1F), SEQ ID NO:14 (RG1G), SEQ ID NO:16 (RG1H), SEQ ID NO:20 (RG1J). Exemplary RG2 polypeptides have the sequences shown in SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

An exemplary RG3 polypeptide has the sequence shown in SEQ ID NO:138. An exemplary RG4 polypeptide has the sequence shown in SEQ ID NO:139. RG polynucleotides will have at least 60% identity, more typically at least 65% identity, generally at least 70% identity, and preferably at least 75% identity, more preferably at least 80% identity, and most preferably at least 85%, 90%, or 95% identity at the deduced amino acid level. The regions where substantial identity is assessed can be inclusive or exclusive of the nucleotide binding site or the leucine rich region.

Vectors and Transcriptional Control Elements

The invention, providing methods and reagents for making novel species and genres of RG nucleic acids described herein, further provides methods and reagents for expressing these nucleic acids using novel expression cassettes, vectors, transgenic
5 plants and animals, using constitutive and inducible transcriptional and translational *cis*-
(*e.g.*, promoters and enhancers) and *trans*-acting control elements.

The expression of natural, recombinant or synthetic plant disease resistance polypeptide-encoding or other (*i.e.*, antisense, ribozyme) nucleic acids can be achieved by operably linking the coding region a promoter (that can be plant-specific or not,
10 constitutive or inducible), incorporating the construct into an expression cassette (such as an expression vector), and introducing the resultant construct into an *in vitro* reaction system or a suitable host cell or organism. Synthetic procedures may also be used.

Typical expression systems contain, in addition to coding or antisense sequence, transcription and translation terminators, polyadenylation sequences, transcription and
15 translation initiation sequences, and promoters useful for transcribing DNA into RNA.

The expression systems optionally at least one independent terminator sequence, sequences permitting replication of the cassette *in vivo*, *e.g.*, plants, eukaryotes, or prokaryotes, or a combination thereof, (*e.g.*, shuttle vectors) and selection markers for the selected expression system, *e.g.*, plant, prokaryotic or eukaryotic systems. To ensure proper
20 polypeptide expression under varying conditions, a polyadenylation region at the 3'-end of the coding region can be included (see Li (1997) *Plant Physiol.* 115:321-325, for a review of the polyadenylation of RNA in plants). The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA (*e.g.*, using *Agrobacterium tumefaciens* T-DNA replacement vectors, see *e.g.*, Thykjaer (1997) *Plant*
25 *Mol Biol.* 35:523-530; using a plasmid containing a gene of interest flanked by *Agrobacterium* T-DNA border repeat sequences; Hansen (1997) "T-strand integration in maize protoplasts after codelivery of a T-DNA substrate and virulence genes," *Proc. Natl. Acad. Sci. USA* 94:11726-11730.

To identify the promoters, the 5' portions of the clones described here are
30 analyzed for sequences characteristic of promoter sequences. For instance, promoter sequence elements include the TATA box consensus sequence (TATAAT), which is usually 20 to 30 base pairs upstream of the transcription start site. In plants, further

upstream from the TATA box, at positions -80 to -100, there is typically a promoter element with a series of adenines surrounding the trinucleotide G (or T) N G (see, *e.g.*, Messing, in *Genetic Engineering in Plants*, pp. 221-227, Kosage, Meredith and Hollaender, eds. 1983). If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the RG coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from viral genes, such as T-DNA.

The nucleic acids of the invention can be expressed in expression cassettes, vectors or viruses which are transiently expressed in cells using, for example, episomal expression systems (*e.g.*, cauliflower mosaic virus (CaMV) viral RNA is generated in the nucleus by transcription of an episomal minichromosome containing supercoiled DNA, Covey (1990) *Proc. Natl. Acad. Sci. USA* 87:1633-1637). Alternatively, coding sequences can be inserted into the host cell genome becoming an integral part of the host chromosomal DNA.

Selection markers can be incorporated into expression cassettes and vectors to confer a selectable phenotype on transformed cells and sequences coding for episomal maintenance and replication such that integration into the host genome is not required. For example, the marker may encode biocide resistance, such as antibiotic resistance, particularly resistance to chloramphenicol, kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or Basta, to permit selection of those cells transformed with the desired DNA sequences, see for example, Blondelet-Rouault (1997) *Gene* 190:315-317; Aubrecht (1997) *J. Pharmacol. Exp. Ther.* 281:992-997. Because selectable marker genes conferring resistance to substrates like neomycin or hygromycin can only be utilized in tissue culture, chemoresistance genes are also used as selectable markers *in vitro* and *in vivo*. See also, Mengiste (1997) "High-efficiency transformation of *Arabidopsis thaliana* with a selectable marker gene regulated by the T-DNA 1' promoter," *Plant J.* 12:945-948, showing that the 1' promoter is an attractive alternative to the cauliflower mosaic virus (CaMV) 35S promoter for the generation of T-DNA insertion lines, the 1' promoter may be especially beneficial for the secondary transformation of transgenic strains containing the 35S promoter to exclude homology-mediated gene silencing.

The endogenous promoters from the RG genes of the present invention can be used to direct expression of the genes. These promoters can also be used to direct expression of heterologous structural genes. The promoters can be used, for example, in recombinant expression cassettes to drive expression of genes conferring resistance to any number of pathogens or pests, including fungi, bacteria, and the like.

Constitutive Promoters

In construction of recombinant expression cassettes, vectors, transgenics, of the invention, a promoter fragment can be employed to direct expression of the desired gene in all tissues of a plant or animal. Promoters that drive expression continuously under physiological conditions are referred to as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include those from viruses which infect plants, such as the cauliflower mosaic virus (CaMV) 35S transcription initiation region; the 1'- or 2'-promoter derived from T-DNA of *Agrobacterium tumefaciens*; the promoter of the tobacco mosaic virus; and, other transcription initiation regions from various plant genes known to those of skill. See also Holtorf (1995) "Comparison of different constitutive and inducible promoters for the overexpression of transgenes in *Arabidopsis thaliana*," *Plant Mol. Biol.* 29:637-646.

Inducible Promoters

Alternatively, a plant promoter may direct expression of the plant disease resistance nucleic acid of the invention under the influence of changing environmental conditions or developmental conditions. Examples of environmental conditions that may effect transcription by inducible promoters include pathogenic attack, anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters. For example, the invention incorporates the drought-inducible promoter of maize (Busk (1997) *supra*); the cold, drought, and high salt inducible promoter from potato (Kirch (1997) *Plant Mol. Biol.* 33:897-909).

Embodiments of the invention also incorporate use of plant promoters which are inducible upon injury or infection to express the invention's plant disease resistance (RG) polypeptides. Various embodiments include use of, *e.g.*, the promoter for a tobacco (*Nicotiana tabacum*) sesquiterpene cyclase gene (EAS4 promoter), which is expressed in wounded leafs, roots, and stem tissues, and upon infection with microbial pathogens (Yin

(1997) *Plant Physiol.* 115(2):437-451); the ORF13 promoter from *Agrobacterium rhizogenes* 8196, which is wound inducible in a limited area adjacent to the wound site (Hansen (1997) *Mol. Gen. Genet.* 254:337-343); the Shpx6b gene promoter, which is a plant peroxidase gene promoter induced by microbial pathogens (demonstrated using a fungal pathogen, see Curtis (1997) *Mol. Plant Microbe Interact.* 10:326-338); the wound-inducible gene promoter *wun1*, derived from potato (Siebertz (1989) *Plant Cell* 1:961-968); the wound-inducible *Agrobacterium pmas* gene (mannopine synthesis gene) promoter (Guevara-Garcia (1993) *Plant J.* 4:495-505).

Alternatively, plant promoters which are inducible upon exposure to plant hormones, such as auxins, are used to express the nucleic acids of the invention. For example, the invention can use the auxin-response elements E1 promoter fragment (AuxREs) in the soybean (*Glycine max* L.) (Liu (1997) *Plant Physiol.* 115:397-407); the auxin-responsive Arabidopsis GST6 promoter (also responsive to salicylic acid and hydrogen peroxide) (Chen (1996) *Plant J.* 10: 955-966); the auxin-inducible *parC* promoter from tobacco (Sakai (1996) 37:906-913); a plant biotin response element (Streit (1997) *Mol. Plant Microbe Interact.* 10:933-937); and, the promoter responsive to the stress hormone abscisic acid (Sheen (1996) *Science* 274:1900-1902).

Plant promoters which are inducible upon exposure to chemicals reagents which can be applied to the plant, such as herbicides or antibiotics, are also used to express the nucleic acids of the invention. For example, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, can be used (De Veylder (1997) *Plant Cell Physiol.* 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. Coding sequence can be under the control of, *e.g.*, a tetracycline-inducible promoter, *e.g.*, as described with transgenic tobacco plants containing the *Avena sativa* L. (oat) arginine decarboxylase gene (Masgrau (1997) *Plant J.* 11:465-473); or, a salicylic acid-responsive element (Stange (1997) *Plant J.* 11:1315-1324. Using chemically- (*e.g.*, hormone- or pesticide-) induced promoters, harvesting of fruits and plant parts would be greatly facilitated. A chemical which can be applied to the transgenic plant in the field and induce expression of a polypeptide of the invention throughout all or most of the plant would make an environmentally safe defoliant or herbicide. Thus, the invention also provides for transgenic plants containing an inducible gene encoding for the RG

polypeptides of the invention whose host range is limited to target plant species, such as weeds or crops before, during or after harvesting.

Abcission promoters are activated upon plant ripening, such as fruit ripening, and are especially useful incorporated in the expression systems (*e.g.*, expression cassettes, vectors) of the invention. In some embodiments, when a plant disease resistant polypeptide-encoding nucleic acid is under the control of such a promoter, rapid cell death, induced by expression of the invention's polypeptide, can accelerate and/or accentuate abscission, increasing the efficiency of the harvesting of fruits or other plant parts, such as cotton, and the like. Induction of rapid cell death at this time would accelerate separation of the fruit from the plant, greatly augmenting harvesting procedures. See, *e.g.*, Kalaitzis (1997) *Plant Physiol.* 113:1303-1308, discussing tomato leaf and flower abscission; Payton (1996) *Plant Mol. Biol.* 31:1227-1231, discussing ethylene receptor expression regulation during fruit ripening, flower senescence and abscission; Koehler (1996) *Plant Mol. Biol.* 31:595-606, discussing the gene promoter for a bean abscission cellulase; Kalaitzis (1995) *Plant Mol. Biol.* 28: 647-656, discussing cloning of a tomato polygalacturonase expressed in abscission; del Campillo (1996) *Plant Physiol.* 111:813-820, discussing pedicel breakstrength and cellulase gene expression during tomato flower abscission.

Tissue-Specific Promoters

Tissue specific promoters are transcriptional control elements that are only active in particular cells or tissues. Plant promoters which are active only in specific tissues or at specific times during plant development are used to express the nucleic acids of the invention. Examples of promoters under developmental control include promoters that initiate transcription only in certain tissues, such as leaves, roots, fruit, seeds, ovules, pollen, pistils, or flowers. Such promoters are referred to as "tissue specific". The operation of a promoter may also vary depending on its location in the genome. Thus, an inducible promoter may become fully or partially constitutive in certain locations.

For example, a seed-specific promoter directs expression in seed tissues. Such promoters may be, for example, ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed coat-specific, or some combination thereof. A leaf-specific promoter has been identified in maize, Busk (1997) *Plant J.* 11:1285-1295. The ORF13 promoter from *Agrobacterium rhizogenes* exhibits high activity in roots (Hansen (1997) *supra*). A maize pollen-specific promoter has been identified in maize (Guerrero (1990)

Mol. Gen. Genet. 224:161-168). A tomato promoter active during fruit ripening, senescence and abscission of leaves and, to a lesser extent, of flowers can be used (Blume (1997) *Plant J.* 12:731-746). A pistil specific promoter has been identified in the potato (*Solanum tuberosum* L.) SK2 gene, encoding a pistil-specific basic endochitinase (Ficker (1997) *Plant Mol. Biol.* 35:425-431). The Blec4 gene from pea (*Pisum sativum* cv. Alaska) is active in epidermal tissue of vegetative and floral shoot apices of transgenic alfalfa, making it a useful tool to target the expression of foreign genes to the epidermal layer of actively growing shoots. The activity of the Blec4 promoter in the epidermis of the shoot apex makes it particularly suitable for genetically engineering defense against insects and diseases that attack the growing shoot apex (Mandaci (1997) *Plant Mol Biol.* 34:961-965).

The invention also provides for use of tissue-specific plant promoters include a promoter from the ovule-specific *BEL1* gene described in Reiser (1995) *Cell* 83:735-742, GenBank No. U39944. Suitable seed specific promoters are derived from the following genes: *MAC1* from maize, Sheridan (1996) *Genetics* 142:1009-1020; *Cat3* from maize, GenBank No. L05934, Abler (1993) *Plant Mol. Biol.* 22:10131-1038; the gene encoding oleosin 18kD from maize, GenBank No. J05212, Lee (1994) *Plant Mol. Biol.* 26:1981-1987; viviparous-1 from *Arabidopsis*, Genbank No. U93215; the gene encoding oleosin from *Arabidopsis*, Genbank No. Z17657; *Atmyc1* from *Arabidopsis*, Urao (1996) *Plant Mol. Biol.* 32:571-576; the 2s seed storage protein gene family from *Arabidopsis*, Conceicao (1994) *Plant* 5:493-505; the gene encoding oleosin 20kD from *Brassica napus*, GenBank No. M63985; *napA* from *Brassica napus*, GenBank No. J02798, Josefsson (1987) *JBL* 26:12196-1301; the napin gene family from *Brassica napus*, Sjobahl (1995) *Planta* 197:264-271; the gene encoding the 2S storage protein from *Brassica napus*, Dasgupta (1993) *Gene* 133:301-302; the genes encoding oleosin a, Genbank No. U09118, and, oleosin B, Genbank No. U09119, from soybean; and, the gene encoding low molecular weight sulphur rich protein from soybean, Choi (1995) *Mol Gen, Genet.* 246:266-268. The tissue specific E8 promoter from tomato is particularly useful for directing gene expression so that a desired gene product is located in fruits. Other suitable promoters include those from genes encoding embryonic storage proteins.

One of skill will recognize that a tissue-specific promoter may drive expression of operably linked sequences in tissues other than the target tissue. Thus, as

used herein a tissue-specific promoter is one that drives expression preferentially in the target tissue, but may also lead to some expression in other tissues as well.

The invention also provides for use of tissue-specific promoters derived from viruses which can include, *e.g.*, the tobamovirus subgenomic promoter (Kumagai
5 (1995) *Proc. Natl. Acad. Sci. USA* 92:1679-1683; the rice tungro bacilliform virus (RTBV), which replicates only in phloem cells in infected rice plants, with its promoter which drives strong phloem-specific reporter gene expression; the cassava vein mosaic virus (CVMV) promoter, with highest activity in vascular elements, in leaf mesophyll cells, and in root tips (Verdaguer (1996) *Plant Mol. Biol.* 31:1129-1139).

10 In some embodiments, the nucleic acid construct will comprise a promoter functional in a specific plant cell, such as in a species of *Lactuca*, operably linked to an RG polynucleotide. Promoters useful in these embodiments include RG promoters. In additional embodiments, the nucleic acid construct will comprise a RG promoter operably linked to a heterologous polynucleotide. The heterologous polynucleotide is chosen to
15 provide a plant with a desired phenotype. For example, the heterologous polynucleotide can be a structural gene which encodes a polypeptide which imparts a desired resistance phenotype. Alternatively, the heterologous polynucleotide may be a regulatory gene which might play a role in transcriptional and/or translational control to suppress, enhance, or otherwise modify the transcription and/or expression of an endogenous gene within the
20 plant. The heterologous polynucleotide of the nucleic acid construct of the present invention can be expressed in either sense or anti-sense orientation as desired. It will be appreciated that control of gene expression in either sense or anti-sense orientation can have a direct impact on the observable plant characteristics.

Modifying and Inhibiting RG Gene Expression

25 The invention also provides for RG nucleic acid sequences which are complementary to the RG polypeptide-encoding sequences of the invention; *i.e.*, antisense RG nucleic acids. Antisense technology can be conveniently used to modify gene expression in plants. To accomplish this, a nucleic acid segment from the desired gene is cloned and operably linked to a promoter such that the anti-sense strand of RNA will be
30 transcribed. The construct is then transformed into plants and the antisense strand of RNA is produced. In plant cells, it has been shown that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, see, *e.g.*,

Sheehy (1988) *Proc. Nat. Acad. Sci. USA* 85:8805-8809; Hiatt et al., U.S. Patent No. 4,801,340.

Antisense sequences are capable of inhibiting the transport, splicing or transcription of RG-encoding genes. The inhibition can be effected through the targeting of genomic DNA or messenger RNA. The transcription or function of targeted nucleic acid can be inhibited, *e.g.*, by hybridization and/or cleavage. One particularly useful set of inhibitors provided by the present invention includes oligonucleotides which are able to either bind RG gene or message, in either case preventing or inhibiting the production or function of RG. The association can be through sequence specific hybridization. Such inhibitory nucleic acid sequences can, for example, be used to completely inhibit a plant disease resistance response. Another useful class of inhibitors includes oligonucleotides which cause inactivation or cleavage of RG message. The oligonucleotide can have enzyme activity which causes such cleavage, such as ribozymes. The oligonucleotide can be chemically modified or conjugated to an enzyme or composition capable of cleaving the complementary nucleic acid. One may screen a pool of many different such oligonucleotides for those with the desired activity.

Antisense Oligonucleotides

The invention provides for with antisense oligonucleotides capable of binding RG message which can inhibit RG activity by targeting mRNA. Strategies for designing antisense oligonucleotides are well described in the scientific and patent literature, and the skilled artisan can design such RG oligonucleotides using the novel reagents of the invention. In some situations, naturally occurring nucleic acids used as antisense oligonucleotides may need to be relatively long (18 to 40 nucleotides) and present at high concentrations. A wide variety of synthetic, non-naturally occurring nucleotide and nucleic acid analogues are known which can address this potential problem. For example, peptide nucleic acids (PNAs) containing non-ionic backbones, such as N-(2-aminoethyl) glycine units can be used. Antisense oligonucleotides having phosphorothioate linkages can also be used, as described in WO 97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144:189-197; Antisense Therapeutics, ed. Agrawal (Humana Press, Totowa, N.J., 1996). Antisense oligonucleotides having synthetic DNA backbone analogues provided by the invention can also include phosphoro-dithioate, methylphosphonate,

phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, and morpholino carbamate nucleic acids, as described herein.

Combinatorial chemistry methodology can be used to create vast numbers of oligonucleotides that can be rapidly screened for specific oligonucleotides that have appropriate binding affinities and specificities toward any target, such as the sense and antisense RG sequences of the invention (for general background information, see, *e.g.*, Gold (1995) *J. of Biol. Chem.* 270:13581-13584).

Inhibitory Ribozymes

The invention provides for with ribozymes capable of binding RG message which can inhibit RG activity by targeting mRNA. Strategies for designing ribozymes and selecting the RG-specific antisense sequence for targeting are well described in the scientific and patent literature, and the skilled artisan can design such RG ribozymes using the novel reagents of the invention. Ribozymes act by binding to a target RNA through the target RNA binding portion of a ribozyme which is held in close proximity to an enzymatic portion of the RNA that cleaves the target RNA. Thus, the ribozyme recognizes and binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cleave and inactivate the target RNA. Cleavage of a target RNA in such a manner will destroy its ability to direct synthesis of an encoded protein if the cleavage occurs in the coding sequence, or, preventing transport of the message from the nucleus to the cytoplasm. After a ribozyme has bound and cleaved its RNA target, it is typically released from that RNA and so can bind and cleave new targets repeatedly.

Catalytic RNA molecules or ribozymes can also be used to inhibit expression of any plant gene. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this cleavage, the ribozyme is not itself altered, and is thus capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs. The design and use of target RNA-specific ribozymes is described, *e.g.*, in Haseioff (1988) *Nature* 334:585-591.

In some circumstances, the enzymatic nature of a ribozyme can be advantageous over other technologies, such as antisense technology (where a nucleic acid

molecule simply binds to a nucleic acid target to block its transcription, translation or association with another molecule) as the effective concentration of ribozyme necessary to effect a therapeutic treatment can be lower than that of an antisense oligonucleotide. This potential advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single
5 ribozyme molecule is able to cleave many molecules of target RNA. In addition, a ribozyme is typically a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding, but also on the mechanism by which the molecule inhibits the expression of the RNA to which it binds. That is, the inhibition is caused by cleavage of the RNA target and so specificity is defined as the ratio
10 of the rate of cleavage of the targeted RNA over the rate of cleavage of non-targeted RNA. This cleavage mechanism is dependent upon factors additional to those involved in base pairing. Thus, the specificity of action of a ribozyme can be greater than that of antisense oligonucleotide binding the same RNA site.

The enzymatic ribozyme RNA molecule can be formed in a hammerhead
15 motif, but may also be formed in the motif of a hairpin, hepatitis delta virus, group I intron or RNaseP-like RNA (in association with an RNA guide sequence). Examples of such hammerhead motifs are described by Rossi (1992) *Aids Research and Human Retroviruses* 8:183; hairpin motifs by Hampel (1989) *Biochemistry* 28:4929, and Hampel (1990) *Nuc. Acids Res.* 18:299; the hepatitis delta virus motif by Perrotta (1992) *Biochemistry* 31:16;
20 the RNaseP motif by Guerrier-Takada (1983) *Cell* 35:849; and the group I intron by Cech U.S. Pat. No. 4,987,071. The recitation of these specific motifs is not intended to be limiting; those skilled in the art will recognize that an enzymatic RNA molecule of this invention has a specific substrate binding site complementary to one or more of the target gene RNA regions, and has nucleotide sequence within or surrounding that substrate
25 binding site which imparts an RNA cleaving activity to the molecule.

Sense Supression

Another method of suppression is sense suppression. Introduction of nucleic acid configured in the sense orientation has been shown to be an effective means by which to block the transcription of target genes. For an example of the use of this method
30 to modulate expression of endogenous genes see, Napoli et al., *The Plant Cell* 2:279-289 (1990), and U.S. Patent No. 5,034,323.

Cloning of RG Polypeptides

Synthesis and/or cloning of RG polynucleotides and isolated nucleic acid constructs of the present invention are provided by methods well known to those of ordinary skill in the art. Generally, the nomenclature and the laboratory procedures in recombinant DNA technology described below are those well known and commonly employed in the art. Standard techniques are used for cloning, DNA and RNA isolation, amplification and purification. Generally enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like are performed according to the manufacturer's specifications. These techniques and various other techniques are generally performed according to Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, (1989).

The isolation of RG genes may be accomplished by a number of techniques. For instance, oligonucleotide probes based on the sequences disclosed here can be used to identify the desired gene in a cDNA or genomic DNA library. To construct genomic libraries, large segments of genomic DNA are generated by random fragmentation, e.g. using restriction endonucleases, and are ligated with vector DNA to form concatemers that can be packaged into the appropriate vector. To prepare a cDNA library, mRNA is isolated from the desired organ, such as roots and a cDNA library which contains the RG gene transcript is prepared from the mRNA. Alternatively, cDNA may be prepared from mRNA extracted from other tissues in which RG genes or homologs are expressed.

The cDNA or genomic library can then be screened using a probe based upon the sequence of a cloned RG gene such as the genes disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species.

Those of skill in the art will appreciate that various degrees of stringency of hybridization can be employed in the assay; and either the hybridization or the wash medium can be stringent. As the conditions for hybridization become more stringent, there must be a greater degree of complementarity between the probe and the target for duplex formation to occur. The degree of stringency can be controlled by temperature, ionic strength, pH and the presence of a partially denaturing solvent such as formamide. For example, the stringency of hybridization is conveniently varied by changing the polarity of the reactant solution through manipulation of the concentration of formamide within the range of 0% to 50%.

Alternatively, the RG nucleic acids of the invention can be amplified from nucleic acid samples using a variety of amplification techniques, such as polymerase chain reaction (PCR) technology, to amplify the sequences of the RG and related genes directly from genomic DNA, from cDNA, from genomic libraries or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

Oligonucleotides can be used to identify and detect additional RG families and RG family species using a variety of hybridization techniques and conditions. Suitable amplification methods include, but are not limited to: polymerase chain reaction, PCR (PCR PROTOCOLS, A GUIDE TO METHODS AND APPLICATIONS, *ed.* Innis, Academic Press, N.Y. (1990) and PCR STRATEGIES (1995), *ed.* Innis, Academic Press, Inc., N.Y. (Innis)), ligase chain reaction (LCR) (Wu (1989) *Genomics* 4:560; Landegren (1988) *Science* 241:1077; Barringer (1990) *Gene* 89:117); transcription amplification (Kwoh (1989) *Proc. Natl. Acad. Sci. USA* 86:1173); and, self-sustained sequence replication (Guatelli (1990) *Proc. Natl. Acad. Sci. USA*, 87:1874); Q Beta replicase amplification and other RNA polymerase mediated techniques (*e.g.*, NASBA, Cangene, Mississauga, Ontario); see Berger (1987) *Methods Enzymol.* 152:307-316, Sambrook, and Ausubel, as well as Mullis (1987) U.S. Patent Nos. 4,683,195 and 4,683,202; Arnheim (1990) *C&EN* 36-47; Lomell *J. Clin. Chem.*, 35:1826 (1989); Van Brunt, *Biotechnology*, 8:291-294 (1990); Wu (1989) *Gene* 4:560; Sooknanan (1995) *Biotechnology* 13:563-564. Methods for cloning *in vitro* amplified nucleic acids are described in Wallace, U.S. Pat. No. 5,426,039.

The degree of complementarity (sequence identity) required for detectable binding will vary in accordance with the stringency of the hybridization medium and/or wash medium. The degree of complementarity will optimally be 100 percent; however, it should be understood that minor sequence variations in the probes and primers may be compensated for by reducing the stringency of the hybridization and/or wash medium as described earlier.

In some preferred embodiments, members of this class of pest resistance genes can be identified by their ability to be amplified by PCR primers based on the sequences disclosed here. Appropriate primers and probes for identifying RG sequences

from plant tissues are generated from comparisons of the sequences provided herein. See, e.g., Table 1. For a general overview of PCR see *PCR Protocols: A Guide to Methods and Applications*. (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), *Academic Press*, San Diego (1990), incorporated herein by reference.

5 Briefly, the first step of each cycle of the PCR involves the separation of the nucleic acid duplex formed by the primer extension. Once the strands are separated, the next step in PCR involves hybridizing the separated strands with primers that flank the target sequence. The primers are then extended to form complementary copies of the target strands. For successful PCR amplification, the primers are designed so that the
10 position at which each primer hybridizes along a duplex sequence is such that an extension product synthesized from one primer, when separated from the template (complement), serves as a template for the extension of the other primer. The cycle of denaturation, hybridization, and extension is repeated as many times as necessary to obtain the desired amount of amplified nucleic acid.

15 In the preferred embodiment of the PCR process, strand separation is achieved by heating the reaction to a sufficiently high temperature for an sufficient time to cause the denaturation of the duplex but not to cause an irreversible denaturation of the polymerase (see U.S. Patent No. 4,965,188). Template-dependent extension of primers in PCR is catalyzed by a polymerizing agent in the presence of adequate amounts of four
20 deoxyribonucleotide triphosphates (typically dATP, dGTP, dCTP, and dTTP) in a reaction medium comprised of the appropriate salts, metal cations, and pH buffering system. Suitable polymerizing agents are enzymes known to catalyze template-dependent DNA synthesis.

Polynucleotides may also be synthesized by well-known techniques as
25 described in the technical literature. See, e.g., Carruthers *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and Adams *et al.*, *J. Am. Chem. Soc.* 105:661 (1983). Double stranded DNA fragments may then be obtained either by synthesizing the complementary strand and annealing the strands together under appropriate conditions, or by adding the complementary strand using DNA polymerase with an appropriate primer
30 sequence.

The present invention further provides isolated RG proteins encoded by the RG polynucleotides disclosed herein. One of skill will recognize that the nucleic acid encoding a functional RG protein need not have a sequence identical to the exemplified genes disclosed here. For example, because of codon degeneracy a large number of nucleic acid sequences can encode the same polypeptide. In addition, the polypeptides encoded by the RG genes, like other proteins, have different domains which perform different functions. Thus, the RG gene sequences need not be full length, so long as the desired functional domain of the protein is expressed.

The resistance proteins are at least 25 amino acid residues in length.

Typically, the RG proteins are at least 50 amino acid residues, generally at least 100, preferably at least 150, more preferably at least 200 amino acids in length. In particularly preferred embodiments, the RG proteins are of sufficient length to provide resistance to pests when expressed in the desired plants. Generally then, the RG proteins will be the length encoded by an RG gene of the present invention. However, those of ordinary skill will appreciate that minor deletions, substitutions, or additions to an RG protein will typically yield a protein with pest resistance characteristics similar or identical to that of the full length sequence. Thus, full-length RG proteins modified by 1, 2, 3, 4, or 5 deletions, substitutions, or additions, generally provide an effective degree of pest resistance relative to the full-length protein.

The RG proteins which provide pest resistance will typically comprise at least one of an LRR or an NBS. Preferably, both are present. LRR and/or NBS regions present in the RG proteins of the present invention can be provided by RG genes of the present invention. In some embodiments, the LRR and/or NBS regions are obtained from other pest resistance genes. See, e.g., Yu *et al.*, *Proc. Natl. Acad. Sci. USA*, 93: 11751-11756 (1996); Bent *et al.*, *Science*, 265: 1856-1860 (1994).

Modified protein chains can also be readily designed utilizing various recombinant DNA techniques well known to those skilled in the art. For example, the chains can vary from the naturally occurring sequence at the primary structure level by amino acid substitutions, additions, deletions, and the like. Modification can also include swapping domains from the proteins of the invention with related domains from other pest resistance genes.

Pests that can be targeted by RG genes and proteins of the present invention include such bacterial pests as *Erwinia carotovora* and *Pseudomonas marginalis*. Fungal pests which can be targeted by the present invention include *Bremia lactucae*, *Marssonina panattoniana*, *Rhizoctonia solani*, *Olpidium brassicae*, root aphid, *Sclerotinia sclerotiorum* and *S. minor*, and *Botrytis cinerea* which causes gray mold. RG genes also provide resistance to viral diseases such as lettuce and turnip mosaic viruses.

Fusion Proteins

RG polypeptides can also be expressed as recombinant proteins with one or more additional polypeptide domains linked thereto to facilitate protein detection, purification, or other applications. Such detection and purification facilitating domains include, but are not limited to, metal chelating peptides such as polyhistidine tracts and histidine-tryptophan modules that allow purification on immobilized metals, protein domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The inclusion of a cleavable linker sequences such as Factor Xa or enterokinase (Invitrogen, San Diego CA) between the purification domain and plant disease resistant polypeptide may be useful to facilitate purification. One such expression vector provides for expression of a fusion protein comprising the sequence encoding a plant disease resistant polypeptide of the invention and nucleic acid sequence encoding six histidine residues followed by thioredoxin and an enterokinase cleavage site (*e.g.*, see Williams (1995) *Biochemistry* 34:1787-1797). The histidine residues facilitate detection and purification while the enterokinase cleavage site provides a means for purifying the desired protein(s) from the remainder of the fusion protein. Technology pertaining to vectors encoding fusion proteins and application of fusion proteins are well described, see *e.g.*, Kroll (1993) *DNA Cell. Biol.*, 12:441-53.

Antibodies Reactive to RG Polypeptides and Immunological Assays

The present invention also provides antibodies which specifically react with RG proteins of the present invention under immunologically reactive conditions. An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. "Immunologically reactive conditions" includes reference to conditions which allow an antibody, generated to a particular epitope of an antigen, to bind to that

epitope to a detectably greater degree than the antibody binds to substantially all other epitopes, generally at least two times above background binding, preferably at least five times above background. Immunologically reactive conditions are dependent upon the format of the antibody binding reaction and typically are those utilized in immunoassay protocols.

"Antibody" includes reference to an immunoglobulin molecule obtained by *in vitro* or *in vivo* generation of the humoral response, and includes both polyclonal and monoclonal antibodies. The term also includes genetically engineered forms such as chimeric antibodies (e.g., humanized murine antibodies), heteroconjugate antibodies (e.g., bispecific antibodies), and recombinant single chain Fv fragments (scFv). The term "antibody" also includes antigen binding forms of antibodies (e.g., Fab', F(ab')₂, Fab, Fv, rIgG, and, inverted IgG). See, Pierce Catalog and Handbook, 1994-1995 (Pierce Chemical Co., Rockford, IL). An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. See, e.g., Huse *et al.* (1989) *Science* 246:1275-1281; and Ward, *et al.* (1989) *Nature* 341:544-546; and Vaughan *et al.* (1996) *Nature Biotechnology*, 14:309-314.

Many methods of making antibodies are known to persons of skill. A number of immunogens are used to produce antibodies specifically reactive to an isolated RG protein of the present invention under immunologically reactive conditions. An isolated recombinant, synthetic, or native RG protein of the present invention is the preferred immunogens (antigen) for the production of monoclonal or polyclonal antibodies.

The RG protein is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies can be generated for subsequent use in immunoassays to measure the presence and quantity of the RG protein. Methods of producing monoclonal or polyclonal antibodies are known to those of skill in the art. See, e.g., Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY; Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY.

Frequently, the RG proteins and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide

variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionucleotides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

The antibodies of the present invention can be used to screen plants for the expression of RG proteins of the present invention. The antibodies of this invention are also used for affinity chromatography in isolating RG protein.

The present invention further provides RG polypeptides that specifically bind, under immunologically reactive conditions, to an antibody generated against a defined immunogen, such as an immunogen consisting of the RG polypeptides of the present invention. Immunogens will generally be at least 10 contiguous amino acids from an RG polypeptide of the present invention. Optionally, immunogens can be from regions exclusive of the NBS and/or LRR regions of the RG polypeptides. Nucleic acids which encode such cross-reactive RG polypeptides are also provided by the present invention. The RG polypeptides can be isolated from any number plants as discussed earlier. Preferred are species from the family *Compositae* and in particular the genus *Lactuca* such as *L. sativa* and such subspecies as *crispa*, *longifolia*, and *asparagina*.

"Specifically binds" includes reference to the preferential association of a ligand, in whole or part, with a particular target molecule (i.e., "binding partner" or "binding moiety") relative to compositions lacking that target molecule. It is, of course, recognized that a certain degree of non-specific interaction may occur between a ligand and a non-target molecule. Nevertheless, specific binding, may be distinguished as mediated through specific recognition of the target molecule. Typically specific binding results in a much stronger association between the ligand and the target molecule than between the ligand and non-target molecule. Specific binding by an antibody to a protein under such conditions requires an antibody that is selected for its specificity for a particular protein. The affinity constant of the antibody binding site for its cognate monovalent antigen is at least 10^7 , usually at least 10^8 , preferably at least 10^9 , more preferably at least 10^{10} , and most preferably at least 10^{11} liters/mole. A variety of immunoassay formats are appropriate for selecting antibodies specifically reactive with a particular protein. For

example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically reactive with a protein. See Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific reactivity. The antibody may be polyclonal but preferably is monoclonal. Generally, antibodies cross-reactive to such proteins as RPS2, RPM1 (bacterial resistances in *Arabidopsis*, L6 (fungal resistance in flax, PRF (resistance to *Pseudomonas syringae* in tomato), and *N*, (virus resistance in tobacco), are removed by immunoabsorption.

Immunoassays in the competitive binding format are typically used for cross-reactivity determinations. For example, an immunogenic RG polypeptide is immobilized to a solid support. Polypeptides added to the assay compete with the binding of the antisera to the immobilized antigen. The ability of the above polypeptides to compete with the binding of the antisera to the immobilized RG polypeptide is compared to the immunogenic RG polypeptide. The percent cross-reactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% cross-reactivity with such proteins as RPS2, RPM1, L6, PRF, and *N*, are selected and pooled. The cross-reacting antibodies are then removed from the pooled antisera by immunoabsorption with these non-RG resistance proteins.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay to compare a second "target" polypeptide to the immunogenic polypeptide. In order to make this comparison, the two polypeptides are each assayed at a wide range of concentrations and the amount of each polypeptide required to inhibit 50% of the binding of the antisera to the immobilized protein is determined using standard techniques. If the amount of the target polypeptide required is less than twice the amount of the immunogenic polypeptide that is required, then the target polypeptide is said to specifically bind to an antibody generated to the immunogenic protein. As a final determination of specificity, the pooled antisera is fully immunosorbed with the immunogenic polypeptide until no binding to the polypeptide used in the immunosorption is detectable. The fully immunosorbed antisera is then tested for reactivity with the test polypeptide. If no reactivity is observed, then the test polypeptide is specifically bound by the antisera elicited by the immunogenic protein.

Production of transgenic plants of the invention

Isolated nucleic acid constructs prepared as described herein can be introduced into plants according techniques known in the art. In some embodiments, the introduced nucleic acid is used to provide RG gene expression and therefore pest resistance in desired plants. In some embodiments, RG promoters are used to drive expression of desired heterologous genes in plants. Finally, in some embodiments, the constructs can be used to suppress expression of a target endogenous gene, including RG genes.

To use isolated RG sequences in the above techniques, recombinant DNA vectors suitable for transformation of plant cells are prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, for example, Weising *et al. Ann. Rev. Genet.* 22:421-477 (1988).

A DNA sequence coding for the desired RG polypeptide, for example a cDNA or a genomic sequence encoding a full length protein, will be used to construct a recombinant expression cassette which can be introduced into the desired plant. An expression cassette will typically comprise the RG polynucleotide operably linked to transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the RG gene in the intended tissues of the transformed plant.

Such DNA constructs may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, PEG poration, particle bombardment and microinjection of plant cell protoplasts or embryogenic callus, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria.

Transformation techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski *et al. Embo J.* 3:2717-2722 (1984).

Electroporation techniques are described in Fromm *et al. Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein *et al. Nature* 327:70-73 (1987).

Agrobacterium tumefaciens-mediated transformation techniques are well described in the scientific literature. See, for example Horsch *et al. Science* 233:496-498 (1984), and Fraley *et al. Proc. Natl. Acad. Sci. USA* 80:4803 (1983). Although *Agrobacterium* is useful primarily in dicots, certain monocots can be transformed by *Agrobacterium*. For instance, *Agrobacterium* transformation of rice is described by Hiei *et al. Plant J.* 6:271-282 (1994). A particularly preferred means of transforming lettuce is described in Micheltore *et al., Plant Cell Reports*, 6:439-442 (1987).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype and thus the desired RG-controlled phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the RG nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans *et al., Protoplast Isolation and Culture, Handbook of Plant Cell Culture*, pp. 124-176, Macmillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al. Ann. Rev. of Plant Phys.* 38:467-486 (1987).

The methods of the present invention are particularly useful for incorporating the RG polynucleotides into transformed plants in ways and under circumstances which are not found naturally. In particular, the RG polypeptides may be expressed at times or in quantities which are not characteristic of natural plants.

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

Detection of RG Resistance Genes

The present invention further provides methods for detecting RG resistance genes in a nucleic acid sample suspected of comprising an RG resistance gene. The means by which the RG resistance gene is detected is not a critical aspect of the invention. For example, RG resistance genes can be detected by the presence of amplicons using RG
5 resistance gene specific primers. Additionally, RG resistance genes can be detected by assaying for specific hybridization of an RG polynucleotide to an RG resistance gene. In some embodiments, the RG resistance gene can be amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide.

In a typical detection method, the nucleic acid sample is contacted with an
10 RG polynucleotide to form a hybridization complex. The hybridization complex may be detected directly (e.g., in Southern or northern blots), or indirectly (e.g., by subsequent primer extension during PCR amplification). The RG polynucleotide hybridizes under stringent conditions to an RG polynucleotide of the invention. Formation of the hybridization complex is directly or indirectly used to indicate the presence of the RG
15 resistance gene in the nucleic acid sample.

Detection of the hybridization complex can be achieved using any number of well known methods. For example, the nucleic acid sample, or a portion thereof, may be assayed by hybridization formats including but not limited to, solution phase, solid phase, mixed phase, or *in situ* hybridization assays. Briefly, in solution (or liquid) phase
20 hybridizations, both the target nucleic acid and the probe or primer are free to interact in the reaction mixture. In solid phase hybridization assays, probes or primers are typically linked to a solid support where they are available for hybridization with target nucleic in solution. In mixed phase, nucleic acid intermediates in solution hybridize to target nucleic acids in solution as well as to a nucleic acid linked to a solid support. In *in situ*
25 hybridization, the target nucleic acid is liberated from its cellular surroundings in such as to be available for hybridization within the cell while preserving the cellular morphology for subsequent interpretation and analysis. The following articles provide an overview of the various hybridization assay formats: Singer *et al.*, *Biotechniques* 4(3):230-250 (1986); Haase *et al.*, *Methods in Virology*, Vol. VII, pp. 189-226 (1984); Wilkinson, "The theory and practice of *in situ* hybridization" In: *In situ Hybridization*, Ed. D.G. Wilkinson. IRL Press, Oxford University Press, Oxford; and *Nucleic Acid Hybridization: A Practical Approach*, Ed. Hames, B.D. and Higgins, S.J., IRL Press (1987).
30

The effect of the modification of RG gene expression can be measured by detection of increases or decreases in mRNA levels using, for instance, Northern blots. In addition, the phenotypic effects of gene expression can be detected by measuring nematode, fungal, bacterial, viral, or other pest resistance in plants. Suitable assays for determining pest resistance are well known. Michelmore and Crute, *Trans. Br. mycol. Soc.*, 79(3): 542-546 (1982).

The means by which hybridization complexes are detected is not a critical aspect of the present invention and can be accomplished by any number of methods currently known or later developed. RG polynucleotides can be labeled by any one of several methods typically used to detect the presence of hybridized nucleic acids. One common method of detection is the use of autoradiography using probes labeled with ^3H , ^{125}I , ^{35}S , ^{14}C , or ^{32}P , or the like. The choice of radioactive isotope depends on research preferences due to ease of synthesis, stability, and half lives of the selected isotopes. Other labels include ligands which bind to antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. Alternatively, probes can be conjugated directly with labels such as fluorophores, chemiluminescent agents or enzymes. The choice of label depends on sensitivity required, ease of conjugation with the probe, stability requirements, and available instrumentation. Labeling the RG polynucleotide is readily achieved such as by the use of labeled PCR primers.

The choice of label dictates the manner in which the label is bound to the probe. Radioactive probes are typically made using commercially available nucleotides containing the desired radioactive isotope. The radioactive nucleotides can be incorporated into probes, for example, by using DNA synthesizers, by nick translation with DNA polymerase I, by tailing radioactive DNA bases to the 3' end of probes with terminal deoxynucleotidyl transferase, by treating single-stranded M13 plasmids having specific inserts with the Klenow fragment of DNA polymerase in the presence of radioactive deoxynucleotides, dNTP, by transcribing from RNA templates using reverse transcriptase in the presence of radioactive deoxynucleotides, dNTP, or by transcribing RNA from vectors containing specific RNA viral promoters (e.g., SP6 promoter) using the corresponding RNA polymerase (e.g., SP6 RNA polymerase) in the presence of radioactive ribonucleotides rNTP.

The probes can be labeled using radioactive nucleotides in which the isotope resides as a part of the nucleotide molecule, or in which the radioactive component is attached to the nucleotide via a terminal hydroxyl group that has been esterified to a radioactive component such as inorganic acids, *e.g.*, ^{32}P phosphate or ^{14}C organic acids, or esterified to provide a linking group to the label. Base analogs having nucleophilic linking groups, such as primary amino groups, can also be linked to a label.

Non-radioactive probes are often labeled by indirect means. For example, a ligand molecule is covalently bound to the probe. The ligand then binds to an anti-ligand molecule which is either inherently detectable or covalently bound to a detectable signal system, such as an enzyme, a fluorophore, or a chemiluminescent compound. Enzymes of interest as labels will primarily be hydrolases, such as phosphatases, esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescers include luciferin, and 2,3-dihydrophthalazinediones, *e.g.*, luminol. Ligands and anti-ligands may be varied widely. Where a ligand has a natural anti-ligand, namely ligands such as biotin, thyroxine, and cortisol, it can be used in conjunction with its labeled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

Probes can also be labeled by direct conjugation with a label. For example, cloned DNA probes have been coupled directly to horseradish peroxidase or alkaline phosphatase, (Renz, M., and Kurz, K. (1984) A Colorimetric Method for DNA Hybridization. *Nucl. Acids Res.* 12: 3435-3444) and synthetic oligonucleotides have been coupled directly with alkaline phosphatase (Jablonski, E., *et al.* (1986) Preparation of Oligodeoxynucleotide-Alkaline Phosphatase Conjugates and Their Use as Hybridization Probes. *Nuc. Acids. Res.* 14: 6115-6128; and Li P., *et al.* (1987) Enzyme-linked Synthetic Oligonucleotide probes: Non-Radioactive Detection of Enterotoxigenic *Escherichia Coli* in Faeca Specimens. *Nucl. Acids Res.* 15:5275-5287).

Definitions

Units, prefixes, and symbols can be denoted in their SI accepted form. Numeric ranges are inclusive of the numbers defining the range. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation, respectively. The

headings provided herein are not limitations of the various aspects or embodiments of the invention which can be had by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully defined by reference to the specification as a whole.

5 As used herein, the term "plant" includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds and plant cells and progeny of same. The class of plants which can be used in the methods of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants.

10 As used herein, "pest" includes, but is not limited to, viruses, fungi, nematodes, insects, and bacteria.

 As used herein, "heterologous" is a nucleic acid that originates from a foreign species, or, if from the same species, is substantially modified from its original form. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form.

15 As used herein, "RG gene," alternatively referred to as "RLG gene," is a gene encoding resistance to plant pests, such as viruses, fungi, nematodes, insects, and bacteria, and which hybridizes under stringent conditions and/or has at least 60% sequence identity at the deduced amino acid level to the exemplified sequences provided herein. RG genes encode "RG polypeptides," alternatively referred to as "RLG polypeptides," which can comprise LRR motifs and/or NBS motifs. The RG polypeptides encoded by RG genes have at least 55% or 60% sequence identity, typically at least 65% sequence identity, preferably at least 70% sequence identity, often at least 75% sequence identity, more preferably at least 80% sequence identity, and most preferably at least 90% sequence identity at the deduced amino acid level relative to the exemplary RG sequences provided herein. The term "RG family" or "RG family genus" or "genus" includes reference to a group of RG polypeptide sequence species that have at least 60% amino acid sequence identity, and, the nucleic acids encoding these polypeptides. The individual species of a genus, i.e., the members of a family, typically are genetically mapped to the same locus.

25 As used herein, "RG polynucleotide" includes reference to a contiguous sequence from an RG gene of at least 18, 20, 25, 30, 40, or 50 nucleotides in length, up to

at least about 100 or at least about 200 nucleotides in length. In some embodiments, the polynucleotide is preferably at least 100 nucleotides in length, more preferably at least 200 nucleotides in length, most preferably at least 500 nucleotides in length. Thus, RG polynucleotide may be a RG gene or a subsequence thereof.

5 As used herein, "isolated," when referring to a molecule or composition, such as, for example, an RG polypeptide or nucleic acid, means that the molecule or composition is separated from at least one other compound, such as a protein, other nucleic acids (*e.g.*, RNAs), or other contaminants with which it is associated *in vivo* or in its naturally occurring state. Thus, an RG polypeptide or nucleic acid is considered isolated
10 when it has been isolated from any other component with which it is naturally associated, *e.g.*, cell membrane, as in a cell extract. An isolated composition can, however, also be substantially pure. An isolated composition can be in a homogeneous state and can be in a dry or an aqueous solution. Purity and homogeneity can be determined, for example, using analytical chemistry techniques such as polyacrylamide gel electrophoresis (SDS-
15 PAGE) or high performance liquid chromatography (HPLC).

The term "nucleic acid" or "nucleic acid molecule" or "nucleic acid sequence" refers to a deoxyribonucleotide or ribonucleotide oligonucleotide in either single- or double-stranded form. The term encompasses nucleic acids, *i.e.*, oligonucleotides, containing known analogues of natural nucleotides which have similar or
20 improved binding properties, for the purposes desired, as the reference nucleic acid. The term also includes nucleic acids which are metabolized in a manner similar to naturally occurring nucleotides or at rates that are improved thereover for the purposes desired. The term also encompasses nucleic-acid-like structures with synthetic backbones. DNA backbone analogues provided by the invention include phosphodiester, phosphorothioate, phosphorodithioate, methylphosphonate, phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, morpholino carbamate, and peptide nucleic acids (PNAs); see *Oligonucleotides and Analogues, a Practical Approach*, edited by F. Eckstein, IRL Press at Oxford University Press (1991); *Antisense Strategies*, Annals of the New York Academy of Sciences, Volume 600, Eds. Baserga and
25 Denhardt (NYAS 1992); Milligan (1993) *J. Med. Chem.* 36:1923-1937; *Antisense Research and Applications* (1993, CRC Press). PNAs contain non-ionic backbones, such as N-(2-aminoethyl) glycine units. Phosphorothioate linkages are described in WO
30

97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144:189-197. Other synthetic backbones encompassed by the term include methyl-phosphonate linkages or alternating methylphosphonate and phosphodiester linkages (Strauss-Soukup (1997) *Biochemistry* 36:8692-8698), and benzylphosphonate linkages (Samstag (1996) *Antisense Nucleic Acid Drug Dev* 6:153-156). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide primer, probe and amplification product. Unless otherwise indicated, a particular nucleic acid sequence includes the complementary sequence thereof.

The term "exogenous nucleic acid" refers to a nucleic acid that has been isolated, synthesized, cloned, ligated, excised in conjunction with another nucleic acid, in a manner that is not found in nature, and/or introduced into and/or expressed in a cell or cellular environment other than or at levels or forms different than the cell or cellular environment in which said nucleic acid or protein is found in nature. The term encompasses both nucleic acids originally obtained from a different organism or cell type than the cell type in which it is expressed, and also nucleic acids that are obtained from the same cell line as the cell line in which it is expressed. invention.

The term "recombinant," when used with reference to a cell, or to the nucleic acid, protein or vector refers to a material, or a material corresponding to the natural or native form of the material, that has been modified by the introduction of a new moiety or alteration of an existing moiety, or is identical thereto but produced or derived from synthetic materials. For example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level, typically, under-expressed or not expressed at all. The term "recombinant means" encompasses all means of expressing, *i.e.*, transcription or translation of, an isolated and/or cloned nucleic acid *in vitro* or *in vivo*. For example, the term "recombinant means" encompasses techniques where a recombinant nucleic acid, such as a cDNA encoding a protein, is inserted into an expression vector, the vector is introduced into a cell and the cell expresses the protein. "Recombinant means" also encompass the ligation of nucleic acids having coding or promoter sequences from different sources into one vector for expression of a fusion protein, constitutive expression of a protein, or inducible expression of a protein, such as the plant disease resistant, or RG. polypeptides of the invention.

The term "specifically hybridizes" refers to a nucleic acid that hybridizes, duplexes or binds to a particular target DNA or RNA sequence. The target sequences can be present in a preparation of total cellular DNA or RNA. Proper annealing conditions depend, for example, upon a nucleic acid's, such as a probe's length, base composition, and the number of mismatches and their position on the probe, and can be readily determined empirically providing the appropriate reagents are available. For discussions of nucleic acid probe design and annealing conditions, see, *e.g.*, Sambrook and Ausubel.

The terms "stringent hybridization," "stringent conditions," or "specific hybridization conditions" refers to conditions under which an oligonucleotide (when used, for example, as a probe or primer) will hybridize to its target subsequence, such as an RG nucleic acid in an expression vector of the invention but not to a non-RG sequence. Stringent conditions are sequence-dependent. Thus, in one set of stringent conditions an oligonucleotide probe will hybridize to only one specie of the genus of RG nucleic acids of the invention. In another set of stringent conditions (less stringent) an oligonucleotide probe will hybridize to all species of the invention's genus but not to non-RG nucleic acids. Longer sequences hybridize specifically at higher temperatures. Stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium (if the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, *i.e.*, about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Often, high stringency wash conditions preceded by low stringency wash conditions to remove background probe signal. An example of medium stringency wash conditions for a duplex of, *e.g.*, more than 100 nucleotides, is 1x SSC at 45°C for 15 minutes (see Sambrook for a description of SSC buffer). An example low stringency wash for a duplex of, *e.g.*, more than 100 nucleotides, is 4-6x SSC at 40°C for 15 minutes. a signal to noise ratio of 2x (or higher) than that observed for an unrelated

probe in the particular hybridization assay indicates detection of a "specific hybridization." Nucleic acids which do not hybridize to each other under stringent conditions can still be substantially identical if the polypeptides which they encode are substantially identical. This can occur, *e.g.*, when a nucleic acid is created that encodes for conservative substitutions. Stringent hybridization and stringent hybridization wash conditions are different under different environmental parameters, such as for Southern and Northern hybridizations. An extensive guide to the hybridization of nucleic acids is found in, *e.g.*, Sambrook, Tijssen (1993) *supra*.

As used herein "operably linked" includes reference to a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame.

In the expression of transgenes one of skill will recognize that the inserted polynucleotide sequence need not be identical and may be "substantially identical" to a sequence of the gene from which it was derived. As explained herein, these variants are specifically covered by this term.

In the case where the inserted polynucleotide sequence is transcribed and translated to produce a functional RG polypeptide, one of skill will recognize that because of codon degeneracy, a number of polynucleotide sequences will encode the same polypeptide. These variants are specifically covered by the term "RG polynucleotide sequence". In addition, the term specifically includes those full length sequences substantially identical (determined as described herein) with an RG gene sequence which encode proteins that retain the function of the RG protein. Thus, in the case of RG genes disclosed here, the term includes variant polynucleotide sequences which have substantial identity with the sequences disclosed here and which encode proteins capable of conferring resistance to nematodes, bacteria, viruses, fungi, insects or other pests on a transgenic plant comprising the sequence.

Two polynucleotides or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence, as described below. The term

"complementary to" is used herein to mean that the complementary sequence is identical to all or a specified contiguous portion of a reference polynucleotide sequence.

The terms "sequence identity," "sequence similarity" and "homology" refer to when two sequences, such as the nucleic acid and amino acid sequences or the polypeptides of the invention, when optimally aligned, as with, for example, the programs PILEUP, BLAST, GAP, FASTA or BESTFIT (see discussion, *supra*). "Percentage amino acid/nucleic acid sequence identity" refers to a comparison of the sequences of two polypeptides/nucleic acids which, when optimally aligned, have approximately the designated percentage of the same amino acids/nucleic acids, respectively. For example, "60% sequence identity" and "60% homology" refer to a comparison of the sequences of two RG nucleic acids or polypeptides which, when optimally aligned, have 60% identity. For example, in one embodiment, nucleic acids encoding RG polypeptides of the invention comprise a sequence with at least 50% nucleic acid sequence identity to SEQ ID NO:1. In other embodiments, the RG polypeptides of the invention are encoded by nucleic acids comprising a sequence with at least 50% sequence identity to SEQ ID NO:1, or, are encoded by nucleic acids comprising SEQ ID NO:1, or, have at least 60% amino acid sequence identity to the polypeptide of SEQ ID NO:2.

"Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 55% or 60% sequence identity, generally at least 65%, preferably at least 70%, often at least 75%, more preferably at least 80% and most preferably at least 90%, compared to a reference sequence using the programs described above (preferably BESTFIT) using standard parameters. One of skill will recognize that these values can be appropriately adjusted to determine corresponding

identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 55% or 60%, preferably at least 70%, more preferably at least 80%, and most preferably at least 95%. Polypeptides having "sequence similarity" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under appropriate conditions. Appropriate conditions can be high or low stringency and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C to about 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent wash conditions are those in which the salt concentration is about 0.02 molar at pH 7 and the temperature is at least about 50°C. However, nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This may occur, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

Nucleic acids of the invention can be identified from a cDNA or genomic library prepared according to standard procedures and the nucleic acids disclosed here used as a probe. Thus, for example, stringent hybridization conditions will typically include at least one low stringency wash using 0.3 molar salt (*e.g.*, 2X SSC) at 65°C. The washes

are preferably followed by one or more subsequent washes using 0.03 molar salt (e.g., 0.2X SSC) at 50°C, usually 60°C, or more usually 65°C. Nucleic acid probes used to identify the nucleic acids are preferably at least 100 nucleotides in length.

As used herein, "nucleotide binding site" or "nucleotide binding domain" ("NBS") includes reference to highly conserved nucleotide-, *i.e.*, ATP/GTP-, binding domains, typically included in the "kinase domain" of kinase polypeptides, such as a kinase-1a, kinase 2, or a kinase 3a motif, as described herein. For example, the tobacco N and Arabidopsis RPS2 genes, among several recently cloned disease-resistance genes, share highly conserved NBS sequence. Kinase NBS subdomains further consist of three subdomain motifs: the P-loop, kinase-2, and kinase-3a subdomains (Yu (1996) *Proc. Acad. Sci. USA* 93:11751-11756). As discussed in detail herein, examples include the Arabidopsis RPP5 gene (Parker (1997) *supra*), the *A. thaliana* RPS2 gene (Mindrinos (1997) *supra*), and the flax L6 rust resistance gene (Lawrence (1995) *supra*) which all encode proteins containing an NBS; and Mindrinos (1994) *Cell* 78:1089-1099; and Shen (1993) *FEBS* 335:380-385. Using the teachings disclosed and incorporated herein and standard nucleic acid hybridization and/or amplification techniques, one of skill can identify members having NBS domains, including any of the genus of NBS-containing plant disease resistant polypeptides of the invention.

As used herein, "leucine rich region" ("LRR") includes reference to a region that has a leucine content of at least 20% leucine or isoleucine, or 30% of the aliphatic residues: leucine, isoleucine, methionine, valine, and phenylalanine, and arranged with approximate repeated periodicity. The length of the repeat may vary in length but is generally about 20 to 30 amino acids. An LRR-containing polypeptide typically will have the canonical 24 amino acid leucine-rich repeat (LRR) sequence, which is present in different proteins that mediates molecular recognition and/or interaction processes; as described in Bent (1994) *Science* 265:1856-1860; Parker (1997) *Plant Cell* 9:879-894; Hong (1997) *Plant Physiol.* 113:1203-1212; Schmitz (1997) *Nucleic Acids Res.* 25:756-763; Hipskind (1996) *Mol. Plant Microbe Interact.* 9:819-825; Tornero (1996) *Plant J.* 10:315-330; Dixon (1996) *Cell* 84:451-459; Jones (1994) *Science* 266:789-793; Lawrence (1995) *Plant Cell* 7:1195-1206; Song (1995) *Science* 270:1804-1806; as discussed in further detail *supra*. Using the teachings disclosed and incorporated herein and standard nucleic acid hybridization and/or amplification techniques, one of skill can

identify polypeptides having LRR domains, including any member of the genus of LRR-containing RG polypeptides of the invention.

The term "promoter" refers to a region or sequence determinants located upstream or downstream from the start of transcription and which are involved in recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating and/or regulating transcription in plant cells; see also discussion on plant promoters, *supra*.

The term "constitutive promoter" refers to a promoter that initiates and helps control transcription in all tissues. Promoters that drive expression continuously under physiological conditions are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation; see also detailed discussion, *supra*.

The term "inducible promoter" refers to a promoter which directs transcription under the influence of changing environmental conditions or developmental conditions. Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters; see also detailed discussion, *supra*.

The term "abscission-induced promoter" or "abscission promoter" refers to a class of promoters which are activated upon plant ripening, such as fruit ripening, and are especially useful incorporated in the expression systems (*e.g.*, expression cassettes, vectors) of the invention. When the plant disease resistant polypeptide-encoding nucleic acid is under the control of an abscission promoter, rapid cell death, induced by expression of the invention's polypeptide, accelerates and/or accentuates abscission of the plant part, increasing the efficiency of the harvesting of fruits or other plant parts, such as cotton, and the like; see also detailed discussion, *supra*.

The term "tissue-specific promoter" refers to a class of transcriptional control elements that are only active in particular cells or tissues. Examples of plant promoters under developmental control include promoters that initiate transcription only (or primarily only) in certain tissues, such as roots, leaves, fruit, ovules, seeds, pollen, pistils, or flowers; see also detailed discussion, *supra*.

As used herein "recombinant" includes reference to a cell, or nucleic acid, or vector, that has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid to a form not native to that cell, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

As used herein, a "recombinant expression cassette" or "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements which permit transcription of a particular nucleic acid in a target cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of the expression vector includes a nucleic acid to be transcribed, and a promoter.

As used herein, "transgenic plant" includes reference to a plant modified by introduction of a heterologous polynucleotide. Generally, the heterologous polynucleotide is an RG structural or regulatory gene or subsequences thereof.

As used herein, "hybridization complex" includes reference to a duplex nucleic acid sequence formed by selective hybridization of two single-stranded nucleic acids with each other.

As used herein, "amplified" includes reference to an increase in the molarity of a specified sequence. Amplification methods include the polymerase chain reaction (PCR), the ligase chain reaction (LCR), the transcription-based amplification system (TAS), the self-sustained sequence replication system (SSR). A wide variety of cloning methods, host cells, and *in vitro* amplification methodologies are well-known to persons of skill.

As used herein, "nucleic acid sample" includes reference to a specimen suspected of comprising RG resistance genes. Such specimens are generally derived, directly or indirectly, from lettuce tissue.

The term "antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments or synthetic or recombinant analogues thereof which specifically bind and recognize analytes and antigens, such as a genus or subgenus of polypeptides of the invention, as described *supra*.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims.

5

EXAMPLES

The following examples are offered to illustrate, but not to limit the claimed invention.

Example 1

10

Example 1 describes the use of PCR to amplify RG genes from lettuce.

Multiple primers with low degeneracy, particularly at the 3' end, were designed based on the sequences of two known resistance genes from tobacco and flax.

DNA Templates

15

Lettuce genomic DNA was extracted from cultivar Diana and a mutant line derived from cultivar Diana using a standard CTAB protocol. To generate cDNA templates, RNA was isolated from cultivar Diana and the mutant following standard procedures; first strand cDNA was synthesized using Superscript reverse transcriptase from 1 Φ g total RNA as specified by the manufacturer (Life Technologies). BAC (bacterial artificial chromosome) clones from the *Dm3* region were isolated from a BAC library of over 53,000 clones using marker AC15 that was known to be closely linked to *Dm3*. Bacterial plasmids containing clones of *L6* and *RPS2* were used as positive controls.

20

PCR with degenerate oligonucleotide primers

25

Oligonucleotide primers were designed based on conserved motifs in the nucleotide binding sites (NBS) of *L6*, *RPS2*, and *N*. Eight primers were made corresponding to the GVGKTT motif in the sense direction; each had 64-fold degeneracy. Six primers were made to the GLPLAL motif in the anti-sense direction; with either 16 or 256-fold degeneracy (Table 1).

30

Oligonucleotides included 14-mer adaptors of (CUA)₄ at the 5' end of the sense primers and (CAU)₄ at the 5' end of the antisense primers to allow rapid cloning of the PCR products into pAMP1 (Life Technologies).

PCR amplification was performed in 50 Φ l reaction volume with 1 Φ M of each of a pair of sense and antisense primers. The templates were denatured by heating to 94EC for 2 min. This was followed by 35 cycles of 30 sec at 94EC, 1 min at 50EC, 2 min at 72EC, with a single final extension of 5 min at 72EC. 25 ng of genomic DNA or cDNA was used. BAC clones as templates required less. The final dNTP concentration was 0.2 mM; $MgCl_2$ was 1.5 mM.

Forty-eight combinations of sense and antisense primers were tested on a panel of nine templates consisting of two genomic DNA samples, two cDNA preparations, three BAC clones and plasmids containing *L6* and *RPS2* as positive controls.

Amplification from *L6* and *RPS2* resulted in fragments of 516 and 513 repectively. Seven combinations of primers resulted in fragments of approximately this size with multiple templates (Table 2). Primers that gave RLG products were: PLOOPAA, PLOOPAG, PLOOPGA, PLOOPGG, PLOOPAC, GLPL3, GLPL4.

(Intentionally left blank)

Table 1

DEGENERATE PRIMER SEQUENCES for NBS PCR

Sense primers based on GVGKTT amino acid sequence from L6, N and rps2 PLOOP motif:

PLOOPAG 5' GGN GTN GGN AAA ACG AC 3'

PLOOPAA 5' GGN GTN GGN AAA ACA AC 3'

PLOOPAT 5' GGN GTN GGN AAA ACT AC 3'

PLOOPAC 5' GGN GTN GGN AAA ACC AC 3'

PLOOPGG 5' GGN GTN GGN AAG ACG AC 3'

PLOOPGA 5' GGN GTN GGN AAG ACA AC 3'

PLOOPGT 5' GGN GTN GGN AAG ACT AC 3'

PLOOPGC 5' GGN GTN GGN AAG ACC AC 3'

Antisense primers based on GLPLAL amino acid sequence:

GLPL1 5' AGN GCN AGN GGN AGG CC 3' .

GLPL2 5' AGN GCN AGN GGN AGA CC 3' !

GLPL3 5' AGN GCN AGN GGN AGT CC 3' '

GLPL4 5' AGN GCN AGN GGN AGC CC 3'

GLPL5 5' AAN GCC AAN GGC AAA CC 3' .

GLPL6 5' AAN GCC AAN GGC AAT CC 3'

TABLE 2. Characteristics of RLGs isolated from lettuce.

	Template	Primers	Number ^a	Size ^b (bp)	Copy number ^c	Dm linkage
5	RLG1	genomic DNA	PLOOPGA+GLPL6	6/6	522	DM4,
		cDNA	PLOOPGA+GLPL6	1/5		DM13
		genomic DNA	PLOOPAA+GLPL6	5/5		
		cDNA	PLOOPAA+GLPL6	1/1		
	RLG2	BACH8	PLOOPGG+GLPL3	3/3	510	DM1, Dm3
	RLG3	genomic DNA	PLOOPGA+GLPL4	3/6	461	Dm5 Dm8
10	RLG4	genomic DNA	PLOOPGA+GLPL4	1/6	524	

^a Number of RLG sequences out of total number of clones sequenced.

^b Size of fragment amplified from the nucleotide binding domain.

^c Estimated copy number from genomic Southern blot analysis and numbers of clones in the BAC library.

Example 2

Example 2 describes the genetic analysis used to obtain a preliminary indication of the linkage relationships of the amplified products and known clusters of resistance genes.

Bulked segregant analysis was performed to obtain a preliminary indication of the linkage relationships of the amplified products and known clusters of resistance genes. DNA from individuals were pooled for each susceptible and resistant bulk. Amplified products were then mapped by RFLP analysis from our intraspecific mapping population. Resistances from four clusters of resistance genes as well as over six hundred markers have now been mapped on this population. Linkage analysis was done using JIONMAP or MAPMAKER mapping programs. Due to a suppression of recombination in the *Dm3* region, sequences were mapped relative to *Dm3* using a panel of deletion mutants that provided greater genetic resolution than the mapping population (Anderson *et al.* 1996). All blots were washed twice at 63EC in 2x SSC/1% SDS for 20 min, followed by one wash at 63EC in 1x SSC/0.1% SDS for 10 or 30 min.

Most of the RLG sequences were analyzed by bulked segregant analysis (BSA) using pools of resistant and susceptible individuals for each of the four clusters of resistance genes. In genomic Southern analyses, all the RLGs revealed numerous fragments of varying intensity. The numbers of bands was highly dependent of the stringency of hybridization. BSA demonstrated that RLG1 was linked to the *Dm4,7* and *Dm13* clusters. Segregation analysis confirmed this linkage.

RLG2 was derived from BAC H8 that was known to be from the *Dm3* region. BSA with RLG2 demonstrated that the polymorphic bands that distinguished the parents of our mapping population mapped to the *Dm1,Dm3* cluster. Several bands absolutely cosegregated with *Dm1* or *Dm3*. To provide finer genetic resolution, RLG2 was also mapped using a panel of *Dm3* deletion mutants. A number of fragments were missing in largest deletion mutant demonstrating that several RLG2 family members are physically located very close to *Dm3*. No fragment was missing in all deletion mutants; however, this is not unexpected as there is extensive duplication within the region.

Example 3

Example 3 describes the screening of a bacterial artificial chromosome library.

Over 53,000 BAC clones containing lettuce genomic DNA were screened with two of the amplified products. High density filters each containing 1536 clones were hybridized to ^{32}P labelled probes. Filters were washed at 65EC with 40 mM $\text{Na}_2\text{PO}_4/0.1\%$ SDS for 5 min followed by 20 min in the same solution.

To isolate additional RLG sequences we screened our genomic BAC library. Clones were identified that hybridized to RLG1 and RLG2. Nearly all the clones that hybridized to RLG2 also hybridized to marker AC15 that had already been shown by deletion mutant analysis to be clustered around *Dm3*. This provided further evidence for clustering of RLG2 sequences.

Using primers conserved within each family, part of the NBS was amplified from each unique BAC clone and sequenced. This revealed that members within each family varied from 64% identical at the deduced amino acid level. The most divergent members only weakly cross-hybridized to each other. Currently, RLG sequences are

considered to be part of the same family of sequences if they are at least 55% identical at the deduced amino acid level and map to the same region of the chromosome.

Example 4:

5 Example 4 describes the cloning, identification, sequencing and characterization of RG polynucleotide sequences; including use of RG sequences from plasmid and PCR products.

 Doubled stranded plasmid DNA clones and PCR products were sequenced using an ABI377 automated sequencer and fluorescently labelled di-deoxy terminators.
10 Sequences were assembled using Sequencher (Genecodes), DNASTar (DNASTar) and Genetics Computer Group (GCG, Madison, WI) software. Database searches were performed using BLASTX and FASTA (GCG) algorithms.

 Sequences flanking the NBS region for RLG2 and for some of RLG1 were obtained by a series of IPCR and the products sequenced directly. IPCR worked less well
15 for RLG1. Therefore RLG1 was subcloned from a BAC clone into pBSK (Stratagene) and the double stranded plasmid sequenced by long range sequencing.

 Initially, a total of 30 clones were sequenced. Three of these seven primer combinations yielded sequences that comprised continuous open reading frames with sequence identity to the NBS of known resistance genes. Seven out of 10 clones amplified
20 from genomic DNA with the primer pair PLOOPGA/GLP6 were 522 bp long; they were identical to each other and named RLG1. All six clones amplified from genomic DNA or cDNA using the primers PLOOPAA/GLP6 were similar/the same as RLG1. All three clones sequenced from BAC clone H8 were 510 bp long, identical to each other but different from RLG1 and were therefore designated RLG2. The 11 clones sequenced from
25 four other primer combinations had no similarity to any NBS motifs and therefore were not studied further. Therefore, sequencing resulted in the identification of clones containing NBS motifs representing four RLG sequences.

 Comparison of the deduced amino acid sequences of RLG1 and RLG2 to those of known resistance genes revealed that RLG1 and RLG2 are as similar to each other
30 as they are to resistance genes from other species and that this is the same level of identity shown between the known resistance genes (Table 3). The percent identity (upper quadrant) and percent identity (lower quadrant) were determined using the MEGALIGN

routine of the DNASTAR package. Identity refers to the proportion of identical amino acids; identity refers to the proportion of identical and similar amino acids and takes into account substitutions of amino acids with similar chemical characteristics. RG1 and RG2 are as similar to each other and to cloned resistance genes as cloned resistance genes from a variety of species are to each other. L6, resistance to *Melampsora lini* in flax (Lawrence *et al.*, 1995). N, resistance to tobacco mosaic virus in tobacco (Whitham *et al.*, 1994). PRF, required for resistance to *Pseudomonas syringae* in tomato. RPS2, resistance to *Pseudomonas syringae* in *Arabidopsis thaliana* (Bent *et al.*, 1994; Mindrinos *et al.*, 1994). RPM1, resistance to *Pseudomonas syringae* pv. *maculicola* in *A. thaliana* (Grant *et al.*, 1995). The initial RG1 and RG2, sequences were amplified from lettuce using degenerate primers.

Table 3**IDENTITIES OF****RESISTANCE GENE HOMOLOGUES**

		RG1	RG2	RG3	RG4	N gene	RPS2
Lettuce	RG1	***	22.7	15.0	29.2	25.4	23.8
Lettuce	RG2		***	32.2	21.6	22.7	33.0
Lettuce	RG3			***	17.2	15.0	32.8
Lettuce	RG4				***	44.3	22.7
Tobacco	N gene					***	21.6
<i>Arabidopsis</i>	RPS2						***

The regions homologous to the primers are included in this analysis as the genomic sequences for RLG1 and RLG2 were determined by IPCR. Interestingly, the genomic sequences for RLG1 exactly matched that of the primers used.

To obtain further evidence that we had amplified resistance genes, we amplified the regions flanking the NBSs of RLG1a and RLG2a by IPCR of BAC clones. These products were then directly sequenced without cloning to minimize the introduction of PCR artifacts. Sequence analysis of the 5' regions failed to detect any homology to known resistance genes. However, the sequence of the 3' region contained leucine-rich

repeats (LRRs). When this sequence was used to search GENBANK using BLASTX, it detected identity to the *Arabidopsis* resistance gene, *RPS2*. This region does not contain as regular LRRs as in some resistance genes; however, the repeat structure seems to be consistent with that of the flax resistance gene, *L6*. Therefore, the presence of an LRR region is further evidence that the sequences we amplified using degenerate oligonucleotide primers are probably resistance genes.

The sequences of the IPCR products also provided the genomic sequences of the regions complementary to the sequences of the degenerate oligonucleotide primers.

The genomic sequences for RLG1 were identical to one of the primers in the mixture.

The RLG sequences are resistance genes as supported by three criteria: the presence of multiple sequence motifs characteristic of resistance genes, genetic cosegregation with known resistance genes, and their existence as clustered multi-gene families. The presence of LRR regions in a similar position relative to the NBS as in cloned resistance genes provides stronger evidence than relying solely sequence similarity between NBS regions.

The clustering of RLG sequences at the same position as the known clusters of resistance genes make them strong candidates for encoding resistance genes. The hybridization patterns and genetic distribution of the RLG sequences are similar to that of cloned resistance genes in other species. Most of these hybridize to small multigene families and preliminary genetic evidence indicates that they are clustered in the genome. Therefore, the degenerate primers that we designed from other resistance genes seemed to have been specific enough to amplify resistance genes rather than P-loop containing proteins in general.

(intentionally left blank)

SEQID NO:1

RLG1A
[Strand]

```
1  ATCGTACCGCTTCGTACGAG  ANCGCTGTCCCTCCTTCATC  TTTTGTGCATATGTCATATTC  TCATNATTTNTGCCACATNT
81  AATTTTGTGGTTATTTTAAA  TTAATTTTATTCCACATGT  CATTTTATGAGTTTTTCTAT  TTTATIGAGTTTCACATAAT
161  ATTCAAATGTAATAACAATA  AATGCATATTTATTTTCTT  TAAATAAACGCATATAATAT  ATAGATTAAATCATATAAT
241  ACATAGGTTAAACTCATATA  ATACATATGTTTCAATCCCG  TTTATTTATATGTCCTCATCC  TTAATTTATTTATTTATTTAT
321  TTATTAGAGTAGATGATCTT  TGTGATATTAAAAATTTAAT  TTGTTCAAAATTTTAAATTA  TTAATAATCCCAATTTGA
401  ATAAAAATTAAAAAAATGGA  CCCACCATTAGTCCATCACT  TTTTCAGCTCATCAATATCG  TGAGTATTCTCCTTCGTTTC
481  CACCTTAATCAATATTTCCA  GCGAATGACAGACTCTACG  GCGTTTCTGAATTTGCGTTC  CGACACTGTTCAATGAAGGA
561  GATAATAAATCAAATGGAGC  TGCTCCAATGTTCAATGCTG  ATGAAAGGTGAATTGTATGT  GAAGANAATGTCAGCGATCN
641  ATCTCCATCCGGAACCCACC  ACATTATCAGTGTACCACCA  AACCACCTCAAACGGYGGAA  GTAGRRKACWRKAAAGTCA
721  TGAAGAATAGATTATTTTTG  TCCTCATGGGCTGACTGAGG  AGCGGGTTTAGTTCATCAT  TTTCTTTGANCAAGAATTA
801  TCGGTCCATCGAATTTTTAC  ATCGACAAAGAAGTTTCACT  TCGCAATGTTTTGTTAAACA  ATTTTAAATCTTTTTATCTT
881  TTCGTTGAAACTCTCAAT  GCAACTTGCACATTTGCAAT  TTGGGGCCCAAAATTTGTG  GTGGGCGTTAAATTTAATCCA
961  CATATTCACTGTAAACAATA  ATTCAAATCGATCTCTGTT  ATCCAATTCATCAACATCTC  TTGATAATTGAAATCATTCA
1041  CGCTTCATCCATTTCTATCCA  CATCTATACTATATTTCTCTG  CTCTTATCATATTAAACGAT  GGCTGAAATCGTTCCTTCTG
1121  CTTCTTTGACAGTGGTGTMT  GAAAAGCTGGCATTTGAAGC  CTGGAAGAAGATTGTTCCGT  CCAAAGAATTGAATCTGAG
1201  CTTAAGAAATTGAAGGAGAC  ATTAGACCAAAATCCAAGAT  TGCTTAACGATGCTTCCAG  AAGGAAGTAACTAATGAAGC
1281  CGTTAAAGATGGCTGAATG  ATCTCCAACATTTGGCTTAT  GACATAGACGACCTACTTGA  TGATTTTGCACAAAGCTG
1361  TTCATCGTGAGTTGACCGAG  GAGGGTGGAGCCTCTCCAG  TATGGTAAGAAAATAATCC  CAAGTGTGTGCACAAGTTTC
1441  TCACAAAGTAATAGGATGCA  TGCCAAGTTAGATGATATTG  CCACCAGGTTACAAGAATCG  GTAGAGCAAAAAATAATCT
1521  TGGTTTAAGTGTGATAACAT  ATGAAAAGCCAAAATTGAA  AGGTATGAGGCGTCTTTGGT  AGATGAAAGCGGTACTGTG
1601  AAGTGAATAAATCATGACAA  AAATTGCTGGGAAGCTGTT  GGGGATAAAGATGAATCAG  GGAGTCAAAACTTCAGCATC
1681  GTGCCCATAGTTGATTTGGG  TGGAGTTGGTAAACAACCTC  TAGCTAGACTTTTGTATGAT  GAAAAGAAAGTGAAGGATCA
1761  CTTGCAACTCAGGCTTGGG  TTTGTGTTTCTGATGAGTTC  AGTGTTCCTAATAAGCAG  AGTTATTTATCAATCTGTGA
1841  CTGGGGAAAGAGAGGAGTTT  GAAGACTTAAATCTGCTTCA  AGAAGCTCTTAAAGAGAAAC  TTAGGAACCAGCTATTTCTA
1921  ATAGTTTTGGATGATGTGTG  GTCTGAAAGCTATGGTGATT  GGGAGAAATTAGTGGGCCCC  TTCCTTGGGGGTCTCCTGG
2001  AAGTGAATAAATCATGACAA  CTCGGAAGGAGCAATTGCTC  CTGGAAGCTGGGCTTTTCTCA  TCAAGACCTCTGGAGGGTC
2081  TATCACAAGATGATGCTTTG  TCTTTGTTTGTCTCAACAGC  ATTTGGTGTACCAAACTTTG  ATTCACATCCAACACTAAGG
2161  CCACATGGAGAACTGTTTGT  GAAGAAATGTGATGGCTTAC  CTCAGCTTTAAGAACACTT  GGAAGGTTATTAAGGACAAA
2241  AACAGACGAGGAACAATGGA  AGGAGCTGTTGGATAGTGAG  ATATGGAGGTTAGGAAAGAG  CGATGAGATTGTTCCGGCTC
2321  TTAGACTAAGCTACAATGAT  CTTTCTGCGCTTTTGAAGC  RTTCTTTGCTATATGCTCCT  TGTTCCTCAAGGACTATGAG
2401  TTTGACAAAGGAGGATTGAT  TCTATTGTGATGGCAGAAAG  GGTTTTTCACCAACCAACT  AYAAACAAGTCAAAGCAAGC
2481  KTTGGGTCTTGAATATTTTR  AAGAGTTRTTGTCAAGTCTC  TTTTTCACATGCTCCTAA  TRRCAAATCTTGTGTTGTA
2561  TGCATGACCTAATGATGAT  TTGGCTACATTTGTTGCTGG  AGAATTTTTTCAAGGTTAG  ACATAGAGATGAAGAAGGAA
2641  TTTAGGATGSAATCTTTGGA  RAAGCACCGCATATGTCAT  TTGTATGTGAGRATTACATA  GGTACAAAARGTTTCGAGCC
2721  ATTTAGAGGAGCTAAAAATT  TGAGAACATTTTATGCAATG  TCTGTGTTGGGTGGTAGAAGA  TTGGAAGATGTTTTACTTAT
2801  CAAACAAGGTCTTGAATGAC  WTACTTCARGATTTACCAT  GTTAAGGGTCTCTRAKTTTGA  TTRRTCTTAYAAATAASYRAG
2881  GTACCARAAKTCGTSGGTAG  TATGAASCACTTGCGGTATC  TTAATCTATCWGRAACTTWA  ATCACMCATTTACCGGAAWA
2961  TKTCTGCAATCTTTATAATT  TACARACCCTGATTGTTCT  GGCTGTGAMTATTTAGTTAA  KTTGCCCAARACCTTCTCAA
3041  ASCTTAAAAATTTGCASCAT  TTTGACATGAGGGRTACTCC  KAAKTTAARAACATGCCCT  TARGGATTGGTGARTTGAAA
3121  ARTCTACAAACTCTCTTMM  TAACATTGGCATAGCAATAA  CCGAGCTTAAGAAGCTTGCA  AAYCTCCATGGGAAARTTTG
3201  TATTGGCGGGCTGGGAAAAA  TGGAAAATGCMGTGGGATGC  ACGTTAAGCGAACTTGTCTC  AAAAAAGGTTWAATGARTTA
3281  NAAACTGGRTTGGGGGTGA  TRAATTTAATGTTTCCGAA  ATGGGAACACTTGA AAAAGA  AGTCTCTCAATGAAGTGATGC
3361  CTCATAATGGTACTCTANAA  AAAACCCANAATTATGTCTA  TAGGGGGTATAGAGTTTCCA  AATTGGGTTGGTTNCACTAA
3441  GGGTTTCTGAAACTAGAGAT  GTGTTTCATGGTGTATGAAA  AGANTGTTTACGTAGTTTC  ATCAATCACCAGTGGGAAA
3521  TAGATGATATTTTCAGGGCY  TACTGATGAGATGTGGAGAG  GTATGATAGGGTNTCTTGGG  GCCGTAGAAGAAATAAGCAT
3601  CCATTCCTGTATGAATAAA  GATATTTGTGGGAATCAGAA  GCAGAGGCAAGTAAGGTTCT  TATGAATTTAAGAAGTTGG
3681  ATTTAGGTGAATGTGAAAAA  TTGGTGAGTTTAGGGGAGAA  AAAGGAGGATAATCTAATA  TTAATAGTGGGAGCAGCCTA
3761  ACATCTTTTAGGAGGTTGAA  TGTATGGAGATGTAAACAGT  TGAGGCAATTCAGGTGTCCA  GATAGCATGGAGAAATTTGTA
3841  TATGCAATGTGTGATTCAA  TNACATCCGTCTCTTCCCA  ACAGGAGGAGGACAGAAAGT  CRAGTCACTTACCATCACTG
3921  ATTGCAAGAAGCTTTCCGAA  GAGGAGTTGGGAGGACGAGA  GAGGACAAGAGTGTCTATAA  ACTCAAAAATGCAGATGCTT
4001  GAATCACTAGATATACGTAA  TTGGCCAAATCTGAAATCTA  TCAGTGAATTGAGTTGCTTC  ATTCACCTGAACAGATTATA
4081  TATATCAAATGTCCGAGTR  TGGAGTCATTTCTGACCAT  GAGTTGCCAAATCTCACCTC  CTTAACAGATCGAAGGAGAG
4161  GACAGCGATTTTCGTACGAA  CGGTTACGATTCGACTGGCC  GTCGTTTT
```

SEQ ID NO: 2

RLG1B
[Strand]

```
1  AACCGTTCGT ACGAGAATCG CTGTCTCTCT CTTCCTGTAA TATAATGATA AGAAAAATA TGATTAAAGG
71  TTTAAATCCA AAATCCATTA TTCCACCGGT GATATGATGC ACTAGCTGTA GTATGCAAAA ACAGTATTAT
141 AAATGCTAAC CAAAACAGCA GCTAAGAAAC AATATAAATA ATGGTTTGAA TCGTCCTTTC TCCGTACAQT
211 CATTTCTTCC AAATCCCTAT CATTCATACA TACAAGTGCT CCCATATTAG GTTTTCACTA TAAGCAATGG
281 CTGAAATCCT TGGTTCCTGG TTCTTTGCGG TGTTCCTTGA AAAGCTTGCT TCTGAAGCCT TGAAGAGGGT
351 TGCTTGCTCC AAAGTAATTG ACAAGGAGCT CGAGAAATTG AATAGCTCAT GAATCAATAT AAAAGCTCTG
421 CTCATATGAT CTCTCTAGAA GGAAATAAGT AAGGAAGCTG TTAAAGAATG GTTGAATGCT CTTCAACATT
491 TGCCTTACGA CATAGATGAT CTAATTGGCG ATTTGGCAAC CAAAGCTATC CATCGTAAGT TCTCTGAGGA
561 ATACGGGGCC ACCATCAACA AGGTACGAAA GTTAATTCCA TCTTGTTTCT CTAGTTTGTG AAGTACTAAG
631 ATGCGCAACA AGATACATAA TATTACCAGC AAGTTACAAG AACTATTAGA AGAGAGAAAT AATCTTGGAT
701 TATGTGAAAT TGGTGAAGC CGAAACTTTC GAAATAGAAA ATCAGAGACC TCINTGCTAG ATCCATCTAG
771 TATTGTTGGA CGCACAGATG ATAAGGAAGC GTTGCTTCTC AAGCTATATG AACCATGTGA TAGAAACTTT
841 AGCATCTTGC CNATAGTTGG TATGGGTGGG TTAGATAAGA CCACTTTAGG TAGACTTTTG TATGATNAAA
911 TGCAAGTGAA GGATCACTTC GAACTCAAGG CGTGGGTTTG TGTTCCTGAT GAGTTTGATA TCTTCGGTAT
981 AAGCAAAACC ATTTTCGAAT CGATAGAGGG GGGAAACCAA GAGTTTAAGG ATTTAAATCT GCTTCAGGTG
1051 GCTTTAAAGG AGAAAATCTC AAAGAAACGA TTTCTTGTG TTCTTGATGA TGTATGGAGC GAGAGCTATA
1121 CTGATTGGGA AATTCTAGAA CGTCCATTTT TAGCAGGAGC ACCAGGAAGT AAAGTAATCA TCACAACCCG
1191 CAAGTTGTCTG TTGCTAAACC AATTGGGTCA TGATCAACCA TACCAATTGT CTGATTTGTC ACATGACAAT
1261 GCTCTATCCT TATTTTGTCA ACACGCATTT GGTGTAAATA GCTTTGATTC ACATCCGATA CTTAAACCAC
1331 ATGGTGAAGG TATGTTGAA AAATGTGATG GTTTGCCATT GGCTTTGATT GCACTTGGGA GGTATTGAG
1401 GACAAAAGA GATGAGGAAG AATGGAAGGA ACTATTGAAT AGTGAGATAT GGAGGTTAGG AAAGAGAGAT
1471 GAGATTATTC CGGTCTTAG ACTAAGCTAT AATGATCTTT CTGCCTCTTT GAAGCAGTTG TTTGCATATT
1541 GCTCCTTGTT CCCCAAAGAC TATGTGTCA ACAAGGAGAA GTTGATTTTA TTATGGATGG CAGAAGGGTT
1611 TTTGCACAA GAAAATACAA ACAAGTCAAT GGAACGCTTA GNTCTTGAAT ATTTTGACGA CTGTGTGTCA
1681 AGGTCATTTT TTCAACATGC ACTCGATGAC AAATCGTTGT TTGTGGTGCA CGACCTCATG AATGACTTGG
1751 CCACATCTGT TGCTGGAGAT TATTTTTTAA GATTAGACAT TGAATGAAA AAGGAAGCTT TGGAAAAATA
1821 CCGCATATG TCATTGTGTT GTGAGAGTTA CATGGTTTAC AAAAGGTTTG AACCATTTAA AGGAGCTAAA
1891 AAATTGAGAA CTTTCTTAGC AATGCCGTGT GGGATGATAA AAAGTTGGAC AACATTTTAC TTATCAAATA
1961 AGGTCTTTGA TGACTTACTT CACGAATTAC CATGTGTTGAG AGTTCTAAGT TTGAGTTATC TTAGCATCAA
2031 GGAGGTACCT GAAATAATAG GCAATTTGAA AACTTGGCG TATCTTAATT TATCACACAC GAGTATCACA
2101 CATTTACCAG AAAATGTCTG CAATCTTTAC AACTTACAAA CATTGATCCT TTGTGGCTGT TGTTTTATAA
2171 CCAAGTTTCC CACAACTTC TTAAAGCTTA GAAATTTACG GCATTTGGAC ATTAGCGATA CTCGCGGTTT
2241 GAAGAAATG TCCTCGGGGA TTGGTGAATT GAAGAACTA CACACYCTCT CCAAGCTCAT TATTGGAGGT
2311 GAAAAAGAC TAAACGAGCT TAAGAACTTA CAAAATCTCC ATG
```

RLG1b - Diana
[Strand]

```
1   TACTACTACT AGAATTCGGT GTTGGAAGA CGACTCTAGC TAGACTTTTG TATGAGGAAA TGCAAGGGAA
71  GGATCACTTC GAACTTAAGG CGTGGGTATG TGTTTCTGAT GAGTTTGATA TCTTCAATAT AAGCAAAATT
141 ATCTTACAAT CGATAGGTGG TGGAAACCAA GAATTTACGG ACTTAAACCT GCTTCGAGTA GCTTTAAAAG
211 AGAAGATcTC AAAGAAAAGa TTTCTTCTTG TTCTTGATGA TGTTTGGAGT GAAAGCTATA CCGATTGGGA
281 AATINTAGAA CCCCCATTTC TTGCAGGGGC ACCTGGAAGT AAGATTATTA TCACCACCCG GAAGCTGTCA
351 TTGTTAAACA AACTCGGTTA CAATCAACCT TACAACCTTT CGGTTTTGTC ACATGAGAAT GCTTTGTCTT
421 TATCTGTGCA GCATGCATTG GGTGAAGATA ACTTCAATTC ACATCCAACA CTTAAACCAC ATGGCGaAGG
491 TATTGTTGAA AAATGTGATG G&TTGCCATT GGCATTGTCTG ACATGATGAT GATG
```

SEQ ID 137

SEQ ID NO: 3

RLGIC
[Strand]

```

1   TCCCGTGCAA CGTNTATCAT TCAGAAGNGC CCAAAGACCA NAGATNTGTT TAANGNTGNT TINTCAGAAGG
71  AAGTAATTGA TGAAGCTGTN AAAAGATGGC TGATTGATNT CCAACAATTG GCTTACGACA CTGANGACNA
141 ACTTGATGAT NTGCGAACAG AAGCTATTCA TCGTGAGTTG ATCCGTGAAA CTGGAGCTTC CNCCAGCATG
211 GTAAGAAAGC TAATCCCAAG TTGTTGCACA AGTTTCTCAC AAAGTAATAG GATGCATGCC AGGTTAGATG
281 ATATTGCCGC TAAGTACAA GAACGGTAG AGGCGAAAA TAATCTTGGT TTAAGTGTGA TAACATACGA
351 AAAACCCAAA ATTGAAAGAG ATGAGGCGTN TTTGGTAGAT GCAAGTGGTA TCATTGGACG TGAAGATGAT
421 AAGAAAAAAT TGCTTCAGAA GCTGTTGGGG GATACTTATG AATCAAGTAG TCAAAACCTC AACATCGTGC
491 CCATAGTTGG TATGGGTGGG GTAGGTAAAA CAACTCTAGC TAGACTTTTG TATGATGAAA AAAAAGTGAA
561 GGATCACTTC GAACTCAGGG TTTGGGTTTG TGTTCCTGAT GAGTTCAGTG TTCCCAATAT AAGCAGAGTT
631 ATCTATCAAT CTGTGACTGG TGAAAACAAA GAATTTGCAG ATTTAAATCT GCTTCAAGAA GCCCTTAAAG
701 AGAAACTTCA GAACAAACTA TTTCTAATAG TTTTAGATGA TGTATGGTCT GAAAGCTATG GTGATTGGGA
771 GAAATTAGTG GGCCCATTTT ATGCTGGGAC TTCTGGAAGT AGAATAATCA TGACTACTCG GAAGGAGCAA
841 TTAAGCAAC AGCTGGGTTT TTCTCATGAA GACCTCTGCG ATAGTATAGA CTCCCTGCAA CGTCTATCAC
911 AAGAAGATGC TTGTCTTTG TTTTCTCAAC ACGCATTTGG TGTACCTAAC TTTGATTAC ATCCAACACT
981 AAGGCCATAT GGGGAACAGT TTGTGAAAA ATGTGGGGGA TTGCCTTTGG CCTTGT

```

SEQ ID NO:4

RLG1D

[Strand]

```
1   CNTACCTTTC TACGAGATCG CTGTCCCTCC TCGATCTGCT TAACGATGCT TCCCAGAAGG AAGTNACTAA
71  TGAAGCOGTT AAAAGATGGC TGAATGATCT CCAACATTTG GCTTATGACA TANACGACCT ACTTGATGAT
141 CTTCACACAS AAAGCTATTG NTCSTGAGTT GACCGANGAA GGTGGAGCCT CCACCAGTAT GGTAAGAAA
211 CTAATCCCAA GTTGTTCAC AAGTTTCTCA CAAAGTTATA GGATGCATGC CAAGTTAGAT GATATTGCCA
281 CCAGGTTACA AGAACTGGTA GAGGCAAAAA ATAATCTTGG TTTAAGTGTG ATAACATATG AAAAGCCCAA
351 AATTGAAAGG TATGAGGCAT CTTTGGTAGA CGAAAGTGGT ATTTTGGAC GTTNAGATGA TNAGAAAAA
421 TTGATGGAGA AGCTGTTGGA GGATAAAGAT GAATCCGGAG TCNAACTTC AGCATCCTGC CCATAATTGG
491 TATGGGTGGA GTTGGCNAAA CAACTCTAGC TAGACTCTTG TTTGATGAAA AGACAGTGAA GGATCACTTC
561 GAACTCAGGG CTGCGGTTTG TGTTCCTGAT GAATTCAGTA TTCTCAACAT AAGCAAAGTT ATCTATCAAT
631 CTGTGACCGG GGAAAAGAAA GAGTTTGAAG ACTTAAATCT GCTTCAAGAA GCTCTTAGAG GGAAACTACA
701 AAACAACTA TTTCTAATAG TTTTGGATGA TGTATGGTCG GAAAGCTATG GTGATTGGGA GAAATTAGTG
771 GGCCCATTTT ATGCTGGGAC TTCTGGAAGT AGAATAATCA TGACTIONCG GAAGGAGCAA TTACTIONAAC
841 AGTTGGGTTT TTCTCATCAA GACCCCTCTG GTTGTATAGA CTCCCTGCAA CGTCTATCAC AAGATGATGC
911 TTTGCTTTTG TTTGCTCAAC ACGCATTTGG TGWCCA
```

RLG1E
[Strand]

```
1  TCTAGCTAGA CTTTTGTATG ACGAGATGCA AGAGAAGGAT CACTTCGAAC TCAAGGCGTG GGTTCGTGTT
71  TCTGATGAGT TTGATATATT CAATATAAGC AAAATTATTT TCCAATCGAT AGGAGGTGGA AACCAAGAAT
141 TTAAGGACTT AAATCTCCTT CAGTAGCTG TAAAAGAGAA GATTTCAAAG AAACGATTTC TACTTGTTC
211 TGATGATGTT TGGAGTGAAA GCTATGCCGA TTGGGAAATT CTGGAACGCC CATTCCTTGC AGGGGCAGCC
281 GGAAGTAAAA TTATCATGAC GACCCGGAAG CAGTCATTGC TAACCAAAC CGGTTACAAG CAACCTTACA
351 ACCTTTCGGT TTTGTCACAT GACAGTGCTC TCTCTTTATT CTGTCAGCAT GCATTGGGTG AAGATAACTT
421 CGATTCAAT CCAACACTTA AACCACATGG CGAAGGCATT GTTGAAAAAT GTGCT
```

SEQ ID NO:5

RLG1F
[Strand]

```
1  ATTTTCNGCT  CNAACAAAN  AAAAGCAATG  GCTGAAATCT  TTCPTTCNGC  ATTCTAGACC  AGTATTCTTT
71  GAAAAGNTGG  CTTCTGAAGC  CTTGAAGAAG  ATCGCTCGCT  TCCATCGGAT  TGATTCTGAG  CTCAGAAAC
141  TGAAGAGGTC  ATTAATCCAG  ATCAGATCTG  TGCTTAATGA  TGCTTCTGAG  AAGGAAATAA  GTGATGAAGC
211  TGTAAAGAA  TGGCTGAATG  GTCTCCAACA  TTTGTCTTAC  GACATAGACG  ACCTACTTGA  TGATTGGCA
281  ACCGAAACTA  TGCATCGTGA  GTTGACCCAC  GGATCTGGAG  CCTCCACCAG  CTTGTAAGAA  AGATAATCCC
351  AACTTGTTC  ACAGATTTCT  CACTAAGTAG  TAAGATGCGT  AACAAAGTTAG  ATAATATTAC  CATCAAGTTA
421  CAAGAACTGG  TAGAGGAAAA  AGATAATCTT  GGCTTAAGTG  TGAAAGGTGA  AAGCCCCAAA  CATACCAACA
491  GAAGATTACA  GACCTCTTTG  GTAGATGCAT  CTAGCATTAT  TGGTCGTGAA  GGTGATAAGG  ATGCATTGCT
561  CCATAAGCTG  CTGGAGGATG  AACCAAGTGA  TAGAACTTT  AGCATCGTGC  CAATAGTTGG  TATGGGTGGT
631  GTGGGTAAGA  CGACTCTAGC  TAGACTTTTG  TATGACGAGA  TGCAAGAGAA  GGATCACTTC  GAACTCAAGG
701  CGTGGGTTTG  TGTTCCTGAT  GAGTTTGATA  TCTTCAATAT  AAGCAAAGTT  ATCTTCCAAT  CGATAGGTGG
771  TGGARACCA  GAATTTAAGG  ACTTAAATCT  CCTTCAAGTA  GCTGTAAAAG  AGAAGATTTT  AAAGAAACGA
841  TTTCTWYTTG  TTCTGGATGA  TGTTTGAGT  GAAAGCTATA  CAGAATGGGA  AATTCTAGCA  CGTCCATTTT
911  TTGCAGGGGC  ACCAGGAAGT  AAGATTATCA  TGACGACCCG  GAAGTTGTCT  TTGCTAACCA  AACTCGGTTA
981  CAATCAACCT  TACAACCTTT  CSGTTTTGTC  ACATGATAAT  GCTTGTCTTT  TATTCTGTCA  GCAYGCATTG
1051  GGTGAAGATA  ACTTCGATTC  ACATCCAACA  CTTAAACCAC  ASGGTGAAAG  TATTGTTGAA  AATGTGACG
1121  GTTTACCATT  GGCTTTTATT  GCACTTGGGA  GRTTGTGAR  GACAAAAACA  GATGAGGAAG  AATGGAARGA
1191  AGTGTGAAT  AGTGAAATAT  GGGGTCAGG  AAAGGGAGAT  GAGATTGTTT  CGGCTCTTAA  ACTAAGCTAC
1261  AATGATCTCT  CTGCTCTTTT  GAAGAAGTTG  TTTGCATACT  GCTCCTTGTT  CCCAAAAGAC  TATGTGTTCT
1331  ATAAGGAGGA  GTTGATTTTG  TTGTGGATGG  CAGAAGGGTT  TTTGCACCAA  TCAACCACAA  GCAAGTCBAT
1401  GGAACGCTTG  GGHCATGAAG  GTTTTGATGA  ATTGTTGTCA  AGATCATTTT  TTCAACATGC  CCCTGATGCC
1471  AAATCGATGT  TTGTGATGCA  TGACCTGATG  AATGACTTGG  CHACATCTGT  TGCTGGAGAT  TTTTTTCAA
1541  GGATGGACAT  TGAGATGAAG  AARGAATTTA  GGAAGGAAGC  TTTGSAAAAG  YAYCGCCATA  TGTCATTTGT
1611  TTGTGAKGAT  TACATGGTKI  ACAAAGGTT  CRAGCCATT  ACAAGGAGCT  AG
```

SEQ ID NO: 6

RLG1G
[Strand]

```
1  GTGAAGGATC ACTTCGAACT CAGGGCTTGG GTTTGTGTTT CTGATGAATT TAATATCCTC AATATAAGCA
71  AAGTAATTTA TCAATCTGTA ACCGGGGAAA AAAAGGAGTT TGAAGACTTA AATCTGCTTC AAGAAGCTCT
141 TAAAGAAAAA CTTTGGAATC AGTTATTTCT AATAGTTC TGATGATGTGT GGTCTGAAAG CTATCGTGAT
211 TGGGAGAAAT TAGTGGGCCC ATTTTTTTCG GGGTCTCCTG GAAGTATGAT TATCATGACA ACTCGGAAGG
281 AGCAATTGCC AAGAAAGCTG GGTTTCCTC ATCAAGACCC TTTGCAAGGT CTATCACATG ACGATGCTTT
351 GTCTTTGTTT GCTCAACACG CATTGGGTGT ACCA
```

SEQ ID NO: 7

RLG1H
[Strand]

```
1  TCTAGCTAGA CTTTGTGATG AGGAAATGCA AGGGAAGGAT CACTTCGAAC TCAAGGCGTG GGTATGTGTT
71  TCTGATGAGT TTGATATCTT CAATATAAGC AAAATTATCT TACAATCGAT AGGTGGTGGA AACCAAGAAT
141 TTACGGACTT AAACCTGCTT CAAGTAGCTT TAAAAGAGAA GATCTCAAAG AAAAGATTTC TTCTTGTTCT
211 TGATGATGTT TGGAGTGAAA GCTATACCGA TTGGGAAATT CTAGAACGCC CATTTCTTGC AGGGGCACCT
281 GGAAGTAAGA TTATTATCAC CACCCGGAAG CTGTCATTGT TAAACAAACT CGGTACAAAT CAACCTTACA
351 ACCTTTCGGT TTTGTCACAT GAGAATGCTT TGTCTTTATT CTGTCAGCAT GCATTGGGTG AAGATAACTT
421 CAATTCACAT CCAACACTTA AACCACATGG CGAAGGTATT GTTGAAAAAT GTGAT
```

SEQ ID NO: 8

RLGI
[Strand]

```
1  TCTAGCTAGA CTTGTGTATG ATGAGATGCA AGAGAAGGAT CACTTTGAAC TCAAGGCGTG GGTATGTGTT
71  TCTGATGAGT TTGATATATT CAATATAAGC AAAATTATTT TCCAATCGAT AGGAGGTGGA AACCAGAAT
141 TTAAGGACTT AAACCTCCTT CAAGTAGCTG TAAAAGAGAA GATTTTAAAG AAACGATTTC TTCTTGTTC
211 TGACGACGTT TGGAGTGAAA GCTATGCCGA TTGGGAAATT NTGGAACGCC CATTTCTTGC AGGGGCAGCC
281 GGAAGTAAAA TTATCATGAC AACCCGAAAG CAGTCATTGC TAACCAAACG CGGTTACAAG CAACCTTACA
351 ACCTTTCCGT TTTGTCACAT GACAGTGCTC TGTCTTTATT CTGTCAGCAT GCATTGGGTG AAGGTAACCT
421 CGATTACAT CCAACACTTA AACCACATGG CGAAGGCATT GTTGAAAAAT GTGCTGGATT GCCATTGGCA
491 TTGTCGACA
```

SEQ ID NO. 9

RLGLJ
[Strand]

```
1  TACTACTACT AGAATTCGGT GTTGGTAAGA CGACTCTAGC TAGACTTTTG TATGAGGAAA TGCAAGGGAA
71  GGATCACTTC GAACCTTAAGG CGTGGGTATG TGTTTCIGAT GAGTTTGATA TCTTCAATAT AAGCAAAATT
141 ATCTTACAAT CGATAGGTGG TGGAAACCAA GAATTTACGG ACTTAAACCT GCTTCGAGTA GCTTTAAAAG
211 AGAAGATCTC AAAGAAAAGa TTCTTTCTTG TTCTTGATGA TGTTTGAGT GAAAGCTATA CCGATTGGGA
281 AATINTAGAA CGCCCATTTT TTGCAGGGGC ACCTGGAAGT AAGATTATTA TCACCACCCG GAAGCTGTCA
351 TTGTAAACA AACTCGGTTA CAATCAACCT TACAACCTTT CGGTTTTGTC ACATGAGAAT GCTTTGTCTT
421 TATTCTGTCA GCATGCATTG GGTGAAGATA ACTTCAATTC ACATCCAACA CTTAAACCAC ATGGCGaAGG
491 TATTGTTGAA AAATGTGATG GaTTGCCATT GGCATTGTCTG ACATGATGAT GATG
```

SEQ ID NO: 10

RLGIA a.a.

IVTVRTR?LSLLHLLSYVIFS?I?PH?ILWLF.INFYSTCHFMSFSILLSFT.YLNVITINAYLFFFK.THIIYR
LKSynt.VKLI.YICSSPVYLYVSSLIYLLFIY.SR.SL.Y.KFNLFKI.NY..SHNLNKIKKNGPTISPSLFQLINIV
SILLRFHPNQYFQRMDSYGVSEFAFRHCSLKEIINQMELLQCSLLMKGELYVK?MSAI?LHPEPTLSV
YHQTTONGGSR?T?KS.RIDYFCPHGLTEERV.FIIFL?KNYRSIEFLHRQRSFTSQCFVKQFLIFLSFR.NS
SIATCNLQLLGPQICGGR.FNPHIHCKQ.FKSISVHPIHQHLLIIEIIHASSISSTSILYSLLLSY.TMAEIVLS
AFLTUVFEKLA?EALKKIVRSKRIESELKKLKETLDQIQDLDNDASQKEVTNEAVKRWLNDLQHLAYDID
DLLDD?ATEAV?RELTEEGGASSSMVRKLIPSCCTSFSQSNRMHAKLDDIATRLQELVEAKNNLGLSVI
TYEKPKEIERYEASLVDESGTVGREDDKKKLEKLLGDKDESGSQNFSIVPIVGMGGVGKTTLARLLYDEK
KVKDHFELRAWVCVSDEFSPNISRVYQSVTGEKKEFEDLNLQEALKEKLRNQLFLIVLDDVWSESY
GDWEKLVGPFLAGSPGSRIMTTRKEQLLRKLGFSHQDPLEGLSQDDALSIFAQHAFGVVPNFDShPTLR
PHGELFVKKCDGLPLALRTLGRLLRTKTDEEQWKELLDSEIWRGKSDEIVPALRLSYNDLSA?LKLLFA
YCSLFPKDYEFDKKEELLLWMAEGFLHQPT?NKSQRGLGLEYF?ELLSRSFFQHAPN?KSLFVMHDLMDND
LATFVAGEFFSRLDIEMKKEFRM?SLEKHRHMSFVCE?YIGYK?FEPFRGAKNLRTFLALSVGVVEDWK
MFYLSNKVLND?LQDLPLLRLV?LI?L?I??VP??VGSM?HLRYLNLS?T?ITHLPE??CNLYNLQTLIV
SGC?YLV?LPKTF?LKNL?HFDNR?TP?LKNMPL?IGELK?LQTLF?NIGIAITELKNL?NLHGK?CIGG
LGKMENAVGCTLSELVSKKV?..NW??G..I.CFPKWEHLKKKSSMK.CLIMVL?KKP?IMSIGGIEFPN
WVGLSRVSETRDVFMVYEK?CFT.FHQSPSGK.MIFSG?TDEMWRGMIG?LGAVEEISIHSCNEIRYLWE
SEAEASKVLMNLKKLDLGECENLVSLGEKKEDNHNINSGLSTSFRRLNVWRCONSLEHCRCPDSMENLY
MHMCDS?TSVSFPTGGGQKIKSLTITDCKLSEEEELGGRETRVLINSKMQLMESVDIRNWPNLKSISEL
SCFIHLNRLYISNCPS?ESFPDHELPNLTSLTDRRRGQRFSYERLRFDWPSF

SEQ ID NO:11

RLGIB a.a.

NRSYENRCPLLPVI...EKI.LKV.IQNPLFHR.YDALAVVCKNSIINANQNSS.ETI.IMV.IVLSPTYTHFFQIPII
HTYKCSHIRFSLAMAEILGSAFFAVFFEKLASEALKRVACSKVIDKELEKLNSS.INIKALLNDASQKEIS
KEAVKEWLNALQHLPYDIDDLLGDLATKAIHRKFSEEYGATINKVRKLIPSCFSSLSSTKMRNKIHNITS
KLQELLEERNLGLCEIGESRKLNRKSETS?LDPSSIVGRITDDKEALLKLYEPCDRNFSILPIVGMGGL
DKTTLGRLLYD?MQVKDHFELKAWWCVSDEFDIFGISKTIFESIEGGNQEFKDLNLLQVALKEKISKKRFL
VLDDVWSESYTDWEILERPFLAGAPGSKVIITTRKLSLLNQLGHDQPYQLSDLSHDNALSIFCQHAFG
VNSFDSHPILKPHGEGIVEKCDGLPLALIALGRLLRTKRDEEEWKELLNSEIWRGKRDEIIP?LRLSYND
LSASLKQLFAYCSLFPKDYVFNKEKLILLWMAEGFLHNENTNKSMERL?LEYFDDLLSRFFQHALDDKS
LFVVDLMNDLATS VAGDYFLRLDIEMKKEALEKYRHMSFVCESYMVYKRFEPFKGAKKLRTFLAMPV
GMIKSWTTTFYLSNKVLDLLHELP LLRVLSLSYLSIKEVPEIIGNLKHRLRYLNL SHTSITHLPENVCNLYN
LQTLILCGCCFITKFPNNFLKLRNLRHLDISDTPGLKKMSSGIGELKNLHTLSKLIIGGENRLNELKNLQNL
H

SEQ ID NO:12

RLG1c aa.

SRAT?IIQK?PKT?D?F????QKEVIDEAVKRWLID?QQLAYDT?D?LDD?ATEAIHRELIRETGAS?S
MVRKLIPSCCTSFSQSNRMHARLDDIAAK?QELVEAKNNLGLSVITYEKP KIERDEA?LVDASGIIIGRED
DKKKLLQKLLGDTYESSQNFNIVPIVGMGGVGKTTLARLLYDEKKVKDHFELRVWVCVSDEFSVPNIS
RVYQSVTGENKEFADLNLQEALKEKLQNKLFILVDDVWSESYGDWEKLVGPFFHAGTSGSRIIMTTR
KEQLLKQLGFSHEDPLHSIDSLQRLSQEDALSLFSQHAFGVPNFDSHPTLRPYGEQFVKKCGGLPLAL

SEQ ID NO:13

RLGID

?T?LRDRCPSSICLTMLPRRK?LMKPLKDG.MISNIWLMT?TTYLMILQ?KAI??ELT?EGGASTSMVRK
LIPSCCTSFSQSYRMHAKLDDIATRLQELVEAKNNLGLSVITYEKPPIERYEASLVDESGIFGR?DD?KK
LMEKLLEDKDESGVKLQHLPIIGMGGVG?TTLARLLFDEKTVKDHFEFRAWVCVSDEFSILNISKVIYQS
VTGEKKEFEDLNLLQEALRGKLQNKLFIVLDDVWSESYGDWEKLVGPFHAGTSGSRIIMTTRKEQLLK
QLGFSHQDPLRCIDSLQRLSQDDALSLFAQHAFG?

SEQ ID NO: 14

RLGIE

LARLLYDEMQEKDHFELKAWVCVSDEFDIFNISKIIFQSIGGGNQEFKDLNLLQVAVKEKISKKRFLLVLD
DVWSESYADWEILERPFLAGAAGSKIIMTTRKQSLCLKGYKQPYNLSVLSHDSALSFCQHALGEDNF
DSHPTLKPHGEGIVEKCA

SEQ ID NO: 15

RLGIF

FSA?NK?KQWLKSFF?HSRPVFFEK?ASEALKKIARFHRIDSELKKLKRSLIQIRSVLNDASEKEISDEA
VKEWLNGLQHLSYDIDDLDDLATETMHRELTTDLEPPPACKKDNPTCCTDFSLSSKMRNKLDNITIKL
QELVEEKDNLGLSVKGESPKHTNRRLQTSVLDASSIIGREGDKDALLHKLLEDEPSDRNFSIVPIVGMGG
VGKTTLARLLYDEMQUEKDHFEKAWVCVSDEFDIFNISKVIFQSIGGG?QEFKDLNLLQVAVKEKISKKR
FL?VLDDVWSESYTEWEILARPFLAGAPGSKIIMTTRKLSLLTKLGYNQPYNLSVLSHDNAISLFCQHA
LGEDNFD SHPTLKP?GESIVEKCDGLPLALIALGRLL?TKTDEEWEVKEVLNSEIWGSGKGDÉIVPALKLS
YNDLSASLKKLFAYCSLFPKDYVFDKEELILLWMAEGFLHQSTTSKSMERLGHEGFDELLSRSFFQHAPD
AKSMFVMHDLMNDLATS VAGDFFSRMDIEMKKEFRKEAL?K?RHMS?VC?DYMV?KRF?P?TRS.

SEQ ID NO: 16

RLG1 G

VKDHFELRAWVCVSDEFNILNISKVIYQSVTGEKKEFEDLNLLQEALKEKLWNQLFLIVLDDVWSESYR
DWEKLVGPFFSGSPGSMIIMTTRKEQLPRKLGFPHQDPLQGLSHDDALS LFAQHAFGVP

SEQ ID NO: 17

RLG 1 H

LARLLYEEMOGKDHFEKAWVCVSDEFDIFNISKIILQSIGGGNQEFDTLNLLQVALKEKISKKRFLLVLD
DWSESYTDWEILERPFLAGAPGSKIITTRKLSLLNKLGYNQPYNLSVLSHENALSFCQHALGEDNFN
SHPTLKPHGEGIVEKCD

SEQ ID NO: 18

RLGI I

LARLVYDEMQEKDHFELKAWVCVSDEFDIFNISKIIFQSIGGGNQEFKDLNLLQVAVKEKILKKRFLLVLD
DVWSESYADWEI?ERPFLAGAAGSKIIMTTRKQSLTKLGYKQPYNLSVLSHDSALSFCQHALGEGNF
DSHPTLKPHGEGIVEKCAGLPLALST

SEQ ID NO: 19

RLG 1J

EFGVGKTTLARLLYEEMQGKDHFEKAWVCVSDEFDIFNISKIILQSIGGGNQEFDTLNLLRVALKEKISK
KRFLVLDDVWSESYTDWEI?ERPFLAGAPGSKIITTRKLSLLNKLGYNQPYNLSVLSHENALSLFCQH
ALGEDNFNSHPTLKPHG?GIVEKCDGLPLALS

SEQ ID NO: 20

SEQ ID NO: 21
RLG 2A

```
1  TTNACACCAT AAATTCCTCNA CCTGNGGGGA CAAAAACCTA AAAATGGTCC ATAATGCNCA AATCAGNAAG
71  GTTGANAAAG CTCCTAAGTTT TTNACCTCCA NCTGATGCNC NVTCTCTNTA AAGTTCCANAT CCAAGCTTGC
141 CCTCCAACTC TANCNCCTTC AATGGCACCT CCTTCTCTTC AAAAGCACAC AAGAACACTT TCAAGCTCAA
211 CCACACTCAC ACAAGCTCTA GAACNAGGGT TAGGGCACAT TTAGGGTTTT GCTCTCTGGA AATGGTGTCT
281 AAAAGTGAGG CCATAATGTT CCTTATATAA GGCTCACACC CACAATTAGG CTTTCAATCT GAACGTANTA
351 CGCCCACTGT AACTATGGT ACGCCCAACG TACTCGGTAG TCTCCGCTC AANAATACAC TCATGAGTAC
421 GCGCAACGTA CTTTCCCTTA CGCCCAAGCT ACTCAAAGC CAACATTCT TTTCAAGGAC TAATTTTGAC
491 AACTTGAGGA AAGAAAAGGA TCAAAGANAT ATACTTGAAT TCCGGGATGT TACAATGAAG TTGANACCTT
561 GGCTAAAAAA TTAAATTGGT TGTGGAAGCC GTTGGCTGAG CAAGCAACAA GGGTAAAAAT CGTAATCTAC
631 AAATGGTGTT ATTTTCTATT TCTTCTTATT ATTTTACTTG ATTTACGGGT AGTTTTTTTT TCTTACAAAA
701 AATATTAAG TTGATAAAGT ATAGCCACTA AAATTGACTT TTCCAAAAAC ATAATGTCAA ATGGTTCGTA
771 TATGTATCAT GTTGATATTAN ATAATGAATA TGATGATNCT GTTCTATTTA ANCCGAAAAA ATTATCTAAT
841 GATTTTATAT TGGAAAACAA AGTTGTGATT TTTNGCATAA TATAATCAAA TCCNCTTTTG TNGGAGGT
911 GGATAAATGT GGTAAATTTA NAACAAGTGT TTTNACNTG AAGGGTNTGG AAAGGTGAA AAAAGTTAAA
981 ATGATAAAT GTTTACACAA ATGTTGTATC CGACTGAATA TATGTTTAA GGATNATTGT ATTAAATTGT
1051 TGATATATAG TAAGCATAAA TATTTAGAAT TGTGACTTAA ATTTATAAGT TATNCAACT GGATTGAAAC
1121 ATTTTGTATA TANATTAGGA ATGAAAATGA GCAACCTTAA CATACTTATC TTTGGTAGTT TGGTTATTAT
1191 ATTTTATTA NAATATAGAA NCATCCCTTT ATTTTAAACC CATATTGTGG ACGGACTTGA ATAAATGGGA
1261 AAAATGTACC TTGCTATTTA GCACAAAAAA ATTATAAAAA TGTACATTGC TATTTAGCAC AAACAAAAAA
1331 AAAAACTTA TCCTTTTTGC ATTAGGTCAC AAAGAAATAT AAAATGGGAA ATGTGTTGCT ATTTAATGCA
1401 CTAAAAGAAA CATTTTTGGC TTTATTAACC CGGGTAAACC AATAGAAAAA TGGAAGTACA TTGTCATTTA
1471 GCATGAAAAA AAATAACTTT CCATTTTTTG CATCCGGTCA CAATAATAGA AAAATGAAAG TACGTTGCTA
1541 TTTAGCGAAA CTAACCTCCT TTTTTCTTTT TGGCATCGTA TCATAAAATA TAGACTAAAA TACGTTAGTT
1611 TTACATTTTT AATACATTGA AATGTCTAAT CCACATGTTA TTCTATAAAA AGGGAAATGT AATTTACTTA
1681 TTCTTTGATT CTTTGGCTTC TTTTGTATAC CCAAAACATC CCTCTATCCA TCTATTCCAA CTAAATAAAT
1751 GAAAACTATA TTCCCTCCAT TGTAGGGATG TTAATAATTT TGTAATTGTT TTTATGCAAA AAAGTGTTTT
1821 TTGTTAACTA GATTAACGAG ATTCATTTTT CAGCATTTTA GGAGAAGTTC ATCCATCTTT TGGATTTGAA
1891 GTGCAAGCCA AGTCTTTTAA CATGGAATAT GAGGTCCCTA TATGCTCAAA AAATAGCAAA TGAGAAATTT
1961 TTTAAATTGG ATCCCATATA AAGAAAATTT GTTAATGGTT GTTTTAATAT TGGTCAATGT GTCCACCGGA
2031 TGAGCAATAT ACTAGTTTAT AAGGGGTAAA GGTGGGTTTG GTGGGCCCAT TTATCTTTAT TATTTCTAAA
2101 AGTCAGAAAT AAGTAAAAAA AATTATAAGA TAATACCAT AAGGATAAAA AATCATTTTA TTTGGACCAA
2171 AGACCAAAGT TGTAAAGGGG CTGTTTGT TT TTTTGTGAA GAGCTGTGCA ACCACTTTTG TCTGCCCGC
2241 ACAGACACAG TGCAGACATA TCCCTCGCA GAGTGTGTGT TTTTGAAG TCCGACAGAC AAAAAACGT
2311 CTGCGCGAGG TCATCTCTGC GCATATATGT GTCACTGTCT TCAAAGGTCT TCAGACCTCA TTTTAACCAA
2381 AAAAAAATAA GACCAACGGT TTTTTTTTTT TTTTNTTCT TTCTCTGTA GCTGAAAAATG CATTTTTAAT
2451 CTTTATGACA TGAAATTAAG TTTGAAAAAT TAATTTATTT CAACAGCTGT AGACGTTAAA AACAAACAGT
2521 CTTCTTGTG CAGACTGTGG ACATTTGGTC CACCTCTTCT ACCGACAGAG CTTGCAGATG TGGTCCGAG
2591 ACTGCAGACA TTTTGGCTTC AAATAAACAA ACATCACCTA ATTTGACTAC ACCACACGGA CCTCCAATGT
2661 AACAAAAAAA AGGTGAAAC AAAGTTGCCT ATTTCTCCAT ATCCAGGGGC CATTTATGTA AGAGTTATCT
2731 AAATTTTAGT TCGGTAGATC AGTTCTCACA TTTTAAACGG GTAAAGTATA TGTGTGTACG CGCGCACCTG
2801 AAAGGTTTGA ANGTAACCTC CAACTGAAN CAANAATCGA TATGAAGTAT CAAGTTAGAG GTTCAATTGG
2871 TGAAGGAATC AGCTGGAGGT TGGGAATCG AGCTTCCACT ATTAAGGTAA AATCCATAAC CCTAAATGTT
2941 GGTACGCTCA TATATCAAAT TCGGTGTTTT GTTGAATGAA AAAAGCATGC TCAAAAAACC AGTGTAAAGC
3011 ACGGTATATG ACATATTTAT AGTTACTGAT AACAAATAT GATAATTTTG GGTTTACGTA AGTTAGGATT
3081 CGTACTTCAA CCAATGTAA TAGTTTTTGT GAGTCTATCT ATGTATTTGG GGAATCACAT TAGCAACGGG
3151 ATTGTACTAG TAATTCGAAA AAGTCTTTTA AATAATTTT CTGTTTATAA TTTATGAATA GTTTAGCGA
3221 CATCTAATAT TAAATAGAAT GTATCTGATA TGAATTAAT GTCCTTAATG TGAACATAGA CCTTTTCCAT
3291 TTAATAATGC CTAATFATTA GTTCTAATC AATAAATTTT AATTTCTGTT TTATGCTTCT AAGACAATAA
3361 AAATCCATGA TTTTACCTTA AATATTAACA AAAATGACCA TAAATAAATA AAAAATTAGG ATACCAAAACC
3431 CCCCCGCCAT GCCCAATGTC TAAATATTCT TGATGCTTTT GCTTTTCCCT CTTTCCCTTG TTAGTCTATT
3501 ATTCTGGAGA GTTTGAGAGA GTTTCATACA AGAAAATTTT AAGAAGAAAG CAAAGGTCCA GGTATCTCT
3571 TTTCTTAATT ATGTATTAAAC TTACAAGCAT TTTTACACG ATCCATGGTT TTTTGTGTAT GTTTTTTCAA
3641 TTGAAACTAG ATTGGGACTT TTGCCCTTGA TGATTCTATA GATATTGCAT GGAGTTGAGA TTGTGTAGA
3711 AAAGTGGTGA ATGAAAAGAG CAAGTGAAGT CATATAGT ATTGGTAATA TATGATGATG AGATAGAGAT
3781 ATGTTAAAC TGGCTAGAAA ATGTTTTTAA TTGAAATTTT AGGTGTGTA ATTTGAAAGA TACCAAGCTA
3851 ATAACCAATT AGTTATGCTA AATAGTTATA AAGAACAACA AACTCGTAGT TTTTTTTTCA TGATTTTCAA
3921 CCTCTTCTGA CCAACTAAA TTATAACAAA ATTGAATATC ATCTCTGCA ATCAATTTTA ACTTTTGTTA
3991 TTATCATCAT GTCTAAAATT GCCACAAGTT TATTTTCATA GTCATATTGG ATTATGAAAG GACTATTTTT
4061 ACCAATTACA TCTTTACTTT ATGCCCAAAG CTAATACAAAT CCGACTAAAC TAAAGGATT TAGGATGCAT
```

SEQ ID NO: 2/
RLG 2A cont.

```
4131 ATAGTTTGCT CCCCATTAT AGATTTCAT CTAATTGTG TATTGTACTA ATTTAGGTGC CACCACAAGT
4201 AAATTCCTGA AATGGATGTC GTTAATGCCA TTCCTAAACC AGTTGTCGAG ACTCTCATGG TACCCGTTAA
4271 GAAACACATA GGGTACCTCA TTCTCTGCAG GCAATATATG AGGGAAATGG GTATCAAAAT GAGGGGATTG
4341 AATGCTACAA GACTTGGTGT CGAAGAGCAC GTGAACCGGA ACATAAGCAA CCAGCTTGAG GTTCCAGCCC
4411 AAGTCAGGGG TTGGTTTGAA GAAAGTAGAA AGATCAATGC AAAAGTGGAA AATTTCCCTA CGCATGTTGG
4481 CAGTTGTTTC AATCTTAAGG TTAGACACGG GGTCGGAAAG AGAGCCTCCA AGATAATTGA GGACATCGAC
4551 AGTGTCAATG GAGAACAATC TATCATCAT TGGAAATGAT ATTCCATTCC TTTAGGAAGA ATTGATTCCA
4621 CGAAAGCATC CACCTCAATA CCATCAACCG ATCATCATGA TGAGTTCCAG TCAAGAGAGC AAACCTTCAC
4691 AGAAGCACTA AACGCACTCG ATCCTAACCA CAAATCCCAC ATGATAGCCT TATGGGGAAT GGGCGGAGTG
4761 GGGAAAGACG CAATGATGCA TCGGCTCAAA AAGGTTGTGA AAGAAAAGAA AATGTTTAAT TTTATAATTG
4831 AGGCGGTTGT AGGGGAAAAA ACAGACCCCA TTGCTATTCA ATCAGCTGTA GCAGATTACC TAGGTATAGA
4901 GCTCAATGAA AAACTAAAC CAGCAAGACT TGAGAAGCTT CGGAAATGGT TTGTGGACAA TTCCTGGTGT
4971 AAGAAGATCC TAGTCATACT CGACGATGTA TGGCAGTTTG TGATCTGAA TGATATTGGT TTAAGTCTCT
5041 TACCAAATCA AGTGTGCGAC TTCAAGGTGT TGTGACATC ACGAGACAAA GATGTTTGCA CTGAGATGGG
5111 AGCTGAAGTT AATTCACACT TTAATGTGAA AATGTTAATA GAAACAGAAG CACAAAGTTT ATTCCACCAA
5181 TTTATAGAAA TTTCCGATGA TGTGTATCCT GAGTCCATA ATATAGGAGT GAATATTGTA AGGAAGTGTG
5251 GGGGTCTACC CATTCGCATA AAAACCATGT CGTGTACTCT TAGAGGAAAA AGCAAGGATG CATGGAAGAA
5321 TGCACTTCTT CGTTTAGAGC ACTATGACAT TGAAAATATT GTTAATGGAG TTTTAAAT TTTTACGAC
5391 AATCTCCAAG ATGAGGAGAC TAAATCCACC TTTTGTCTTT GTGGAATGTA TCCCGAARAC TTTGATATTC
5461 TTACCGAGGA GTTGTGAGG TATGGATGGG GGTGAAATTT ATTTAAAAAA NIGTATACTA TAGGAGAAGC
5531 AAGAACCAGG CTCACACAT GCATTGAGCG GCTCATTCAT ACAAATTTGT TGATGGAAGT TGATGATGTT
5601 AGGTGCATCA AGATGCATGA TCTGTCTCGT GCTTTTGTMT TGGATATGTA TTCTAAAGT TACTGCTT
5671 CCATTGTCAA CCATAGTAAT ACACTAGAGT GGCATGCAGA TAATATGCAC GACTCTTGTA AAAGACTTTC
5741 ATTAACATGC AAGGGTATGT CTAAGTTTCC TACAGACCTG AAGTTTCCAA ACCTCTCCAT TTTGAAACTT
5811 ATGCATGAAG ATATATCATT GAGGTTTCCC AAAAATCTTT ATGAAGAAAT GGAGAAGCTT GAGGTATATAT
5881 CCTATGATAA AATGAAATAT CCATTGCTTC CCTCATCC TCATGTTCC GTCACCTTC CGGTGTTTCA
5951 TCTACATAAA TGCTCGTTAG TGATGTTTGA TGTCTCTTGT ATTGGAATC TGTGGAATCT AGAAGTCTCT
6021 AGCTTTGCTG ATTCTGCCAT TGACCGGTTG CTTTCCACAA TCGGAAAGTT GAAGAAGCTA AGGCTACTGG
6091 ATTTGACGAA TTGTTATGGT GTTCGTATAG ATAATGTTGT CTTAAAAAAA TTGGTCAAAC TGGAGGAGCT
6161 CTATATGACA GTGGTTGATC GAGGTCGAAA GCGGATTAGC CTCACAGATG ATAAGTCAA GGAGATGGCA
6231 GAGCGTTCAA AAGATATTTA TGCAATAGAA CTTGAGTTCT TTGAAAACGA TGTCAACCA TGTCAACCA
6301 CATTTGAGAA GCTACAACGA TTCCAGATCT CAGTGGGCG CTATTTATAT GGAGATTCCA TAAAGAGTAG
6371 GCACTCGTAT GAAAACACAT TGAAGTTGGT TCTTGA AAAA GGTAATTAT TGGAAGCTCG AATGAACGAG
6441 TTGTTTAAAG AACAGAGGT GTTATGTTTA AGTGTGGGAG ATATGAATGA TCTTGAAGAT ATTGAGGTTA
6511 AGCTATCTCT ACAACTTCTT CAATCTCTTT CGTTCAACAA TTTAAGAGTC CTGTCTGTTT CAAAGTGTGC
6581 AGAGTTGAAA CACTTCTTCA CACCTGGTGT TGCAAACT TTA AAAAAGC TTGAGCATCT TGAAGTTTAC
6651 AAATGTGATA ATATGGAAGA ACTCATACGT AGCAGGGGTA GTGAAGAAGA GACGATTACA TTCCCAAGC
6721 TGAAGTTTTT ATCTTTGTGT GGGCTACCAA AGCTATCCGG TTTGTGCGAT AATGTCAAAA TAATTGAGCT
6791 ACCACAACCT ATGGAGTTGG AACTTGACGA CATTCAGGT TTCACAAGCA TATATCCCAT GAAAAAGTTT
6861 GAAACATTTA GTTTGTGAA GGAAGAGTTA TATATAAAT TTTAATGCTA ATACATTACA AAGGATCTTT
6931 TCAGTTAAAT CTTTCAAAAT ATATTGTAAT TTGATTGTAT GGGGTATTAT TGTGGATGG GACTATTAAAT
7001 AAATGATTAT CTGTCAGGTT CTGATTCTTA AGTTAGAGAA ACTGCATGTT AGTAGTATGT GGAATCTGAA
7071 GGAGATATGG CCTTGCGAAT TTAATATGAG TGAGGAAGTT AAGTTCAAGAG AGATTAAAGT GAGTAACGTG
7141 GATAAGCTTG TGAATTTGTT TCCGCACAAG CCCATATCTC TGCTGCATCA TCTTGAAGAG CTTAAAGTCA
7211 AGAATTGTGG TTCCATTGAA TCGTTATTCA ACATCCATTT GGATTGTGTT GGTGCAACTG GAGATGAATA
7281 CAACAACAGT GGTGTAAGAA TTATTAAAGT GATCAGTTGT GATAAGCTTG TGAATCTCTT TCCACACAAAT
7351 CCCATGTCTA TACTGCATCA TCTTGAAGAG CTTGAAGTGG AGAATTGTGG TTCCATTGAA TCGTTATTCA
7421 ACATTGACTT GGATTGTGCT GGTGCAATTG GGCAAGAAGA CAACAGCATC AGCTTAAAGAA ACATCAAAGT
7491 GGAGAATTTA GGAAGCTAA GANAGGTGTG GAGGATAAAA GGTGGAGATA ACTCTCGTCC CCTTGTTCAT
7561 GGCTTTCAAT CTGTGAAAG CATAAGGGTT GAGNAATGTN AGAAGTTT TAG AATGTATTT ACACCTACCA
7631 CCACAAATTT TAATCTGGGG GCACTTTTGG AGATTTCAAT AGATGACTGC GGAGAAAACA GGGGAAATGA
7701 CGAATCGGAA GAGAGTAGCC ATGAGCAAGA CGAGGTAAGG ATTTCAATTT CACTGTCTTA ATTAATGATT
7771 AAGCTCCTGC TTTTGAATA AAAAAGGGAC AAACCATTTT ATGACTTAAT GTAGCAATAC AAGTCATGTA
7841 TAAGAGTGAC CAACTCTTTT TTATTTATAA AATGACTACA AAATATTTT TTTTATTAGA GATCATGTAT
7911 AAATGTGACT AATTTTTCAT CACCTAACCT TAATGTATAA ATCTTTATAA ATGTCACTAG TTTACTTTTCA
7981 GTAAAATAAC AAATTTAATA AATTATCAAC AAAAAGCATC AACTAAAAAA ATCCCAAC CCGTAATAAT
8051 TTAATAATAA AGGATTTAAC ATCTAATACG AACAATTTTT TTTCTAAACA TGATTGGAC CAAATATCAC
8121 CAGCAACTCA AGTTTGAAT CGATTGAGCT TAAAACCTGA CCAGCATAAT TAGATAGATG AGAGTTGAAG
8191 CTAAAGTGCC TATATAAGTT CGTTTCATCT TTTTCTTGA TCTTGATAGC AAGTTGAATG ATTTTCTTCT
```

RLG 2A cont.

8261 TCAAAATTGA TAAAAATCTA CATTATAAAG AGACTAGCTT GAAAAAAAAT GGTCTAGGTG GGTCTTGGGT
8331 TCTGGTAGAT GAAGATGGAA GGGGAGAGTA TGATTTCAAA GACACAACAC ATCCTTCATT TTATTTATTT
8401 ATTATTATTA TTATTTTTTG ATATCTTGCT CATATTTGTT ACAGATATGT GAGGTCTATT AATCTTTTETA
8471 AATATATAAA AAAATAAATA ACATAAATGA GAAAATTAAA TAAAGAATAA ATTAATAAGG GCACAATAGT
8541 CTTTTTAGGT AAGACAAGGA CCAAACACGC AACAAAATA AACAGTAGGG ACCATCCGAT TTAACAAAAA
8611 TAATTAGGGA CCAAAAACAT AAATTCCCCC AAACCATAGG GACCATTCTT GTAATTTACT CTTACTTTTC
8681 GTTTTGTTC TATTTGGGTA ACTATTTTTT TTGTACACAT CTAGGTAACG AACTTGTGTA AGTGTTCCTA
8751 TTTAGGATGT GACCTACTAC AACCGATCAT AATAGTCATA TGTGAACACT TCCAACAAC TTTATTACTTA
8821 GGTGTGTACA AAAAAACAAT AGTTACCATG ATGTGAACAT ACTGAAAAAT TAATTACCTT AGCAAGTTAT
8891 TTTCCCATTT AGGTGTGATG GAAACAGTTC CGTGAGACCG TGACTTGGAT GGTAGATAAA TTTAGTAAAC
8961 TTAACCCCTC AATTAACCTA CCTTTTTCTT ATTAACCTAA TTTCAACCTA AATTCTGATT CTTGTTTGAA
9031 AGTAAGTTGC ATCTTTATTT TGTATTATC TTGTTGCATA GGATCCTTAG CATCTTTTAA TAATTTATTT
9101 GAAGGTGAAA GATCCAACCTA TTTTAACTCT GTTGGCATTT TCCATCATTT GCAACTGTTT CTTGAAAAAA
9171 AAATACCTAA AATCAAAATA ACCATTTTCA AATCCAAAAT TATAAGAGAG AATTGTAAAT GGACATGGAA
9241 TCATAAATCA TTAACACAGT TCAGTAAACA AGTTGCTAAT TACATTTCTT GCTGTGCAGA TTGAAATTTCT
9311 ATCAGAGAAA GAGACATTAC AAGAAGCCAC TGACAGTATT TCTAATGTTG TATTTCCATC CTGTCTCATG
9381 CACTCTTTTC ATAACCTCCA GAAACTTATA TTGAACAGAG TTAAGGAGT GGAGGTGGTG TTTGAGATAG
9451 AGAGTGAGAG TCCAACAAGT AGAGAATTGG TAACAACCTA CCATAACCAA CAACAACCTA TTATACTTCC
9521 CAACCTCCAG GAATTGATTC TATGGAATAT GGACAACATG AGTCATGTGT GGAAGTGCAG CAACTGGAAT
9591 AAATCTCTCA CTCTTCCAAA ACAACAATCA GAATCCCCAT TCCACAACCT CACAACCATA AAAATTATGT
9661 ATTGCAAAAG CATTAAGTAC TTGTTTTCCG CTCTCATGGC AGAACTTCTT TCCAACCTAA AGCATATCAA
9731 GATAAGAGAG TGTGATGGTA TTGGAGAAGT TGTTTCAAAC AGAGATGATG AGGATGAAGA AATGACTACA
9801 TTTACATCTA CCCACACAAC CACCACTTTG TTCCCTAGTC TTGATTCTCT CACTCTAAGT TTCTTGAGA
9871 ATCTGAAGTG TATTTGGTGA GGTGGTGCCA AGGATGAAGG GAGCAATGAA ATATCTTTCA ATAATACCAC
9941 TGCAACTACT GCTGTCTTGG ATCAATTTGA GGTATGCTTT GTACATATTC AATTATTTAT TTAATTTCTT
10011 TTTTTATTTG CAATATTCTA TAAATAATAC ATTTTATACC CACTATACTA AGATAATAAT TACCTAGAGG
10081 GATGGATGCT ATGACACAGC TGCTACACTT CAGAACTCT AGTAAGGGCA GTTATGGAAG TTCAATAAAA
10151 TGATAATGGC ATCTTTTGAT GGGTAATATA GGCAATTTAA GTTTTATTTT TGTAAAGCA GTATTTAGCA
10221 AGTACTGGCC AGTAGGAGAG GAGAATATCA CCTTTTGTTG AAATCTGGTC ATGTACCCA GAATTTAGTT
10291 AAATGTAACA TTTTAGATAT CAGGGGTCAT CAGGTGACAG ATATTGTAGA ATAGAACAAT ATATAATATC
10361 ACCCAAACT ATTTTCTCTA AGGTATTCTT GTTAAATATG TGCTTTCTTG TTTTCATNGA ATNGCATTC
10431 GTATAATTTA GGTGTTAAAG TGATTTTNTC TTCAATAAAT CCCGAAATTA ATTAAAAAAA AAAAAACAAA
10501 AGTACATTTT TGATGTGGAG AGCACTGGTA TCACTTAGTA TATAAAAAGC TTGATTTTGA ATTAACCTTC
10571 TTATACAAA GTTGTGTATA TAGTTTAATT AGTTTACAT CATTTTCCA TGTGGTGTG CAGTTGCTG
10641 AAGCAAGTGG TGTTCCTTGG AGCTTATGCC AATACGCTAG AGAGATGAGA ATAGAATTCT GCAATGCATT
10711 GTCAAGTGTA ATCCATGTT ATGCAGCAGG ACAAATGCAA AAGCTGAAGG AGAGGACAGC GATTCTCGTA
10781 CGAACGGTTA CGATTGACT GGCCGTCGTT TTACA

SEQ ID NO: 21

RLG1A a.a.

MDVVNAILKPVVETLMVPVKKHIGYLISCRQYMREMGIKMRGLNATRLGVEEHVNRNISNQLEVPAAQV
RGWFEEVVGKINAKVENFPSDVGSCFNLKVRHGVGKRASKIIEDIDSVMREHSIIIWNDHSIPLGRIDSTK
ASTSIPSTDHHDEFQSREQTFTEALNALDPNHKSHMIALWGMGGVGKTTMMHRLKKVVKEKKMFNFII
EAVVGEKTDPIAIQSAVADYLGIELNEKTKPARTEKLRKWFVDNSGGKKILVILDDVWQFVDLNDIGLS
PLPNQGVDFKVLTSRDKDVCTEMGAENVSTFNVKMLIETEAQSLFHQFIEISDDVDPELHNIGVNIVRK
CGGLPIAIKTMACTLRGKSKDAWKNAALLRLEHYDIENIVNGVFKMSYDNLQDEETKSTFLLCGMYPE?FD
ILTEELVRYGWGLKLFKK?YTIGEARTRLNTCIERLIHTNLMEVDDVRCIKMHDLVRAFLDMYSKVEH
ASIVNHSNTLEWHADNMHDSCKRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFPKNFYEEMEKELE
VISYDKMKYPLLPSSPQCSVNLRVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSAIDRLPSTIGKLKCLR
LLDLTNCYGVRIIDNGVLKLVKLEELYMTVVDGRKAISLTDDNCKEMAERSKDIYALELEFFENDAQPK
NMSFEKLQRFQISVGRYLYGDSIKSRHSYENTLKLVEKGELLEARMNELFKKTEVLCLSVGDMNDLEDIE
VKSSSQLLQSSSFNNLRVLVWSKCAELKHFFTPGVANTLKKLEHLEVYKCDNMEELIRSRGSEEETITFP
KLKFLSLCGLPKLSGLCDNVKIIELPQLMELELDDIPGFTSIYPMKKFETFSLLEEVLIPKLEKLHVSSM
WNLKEIWPCEFMSEEVKFREIKVSNCDKLVNLFPHKPISELLHLEELKVKNCGSIESLFNIHLDCVGAT
GDEYNNSGVRIIKVISCDKLVNLFPHNPMSILHHLEEELEVENCGSIESLFNIDLDCAGAIGQEDNSISLRNI
KVENLGKLR?VWRIKGGDNSRPLVHGFQSVESIRVTKC?KFRNVFTPTTTNFNLGALLEISIDDCGENR
GNDESEESSHEQEIEILSEKETLQEATDSISNVVFPSCLMHSFHNLOKLILNRVKGVEVVFIESESPTS
RELVTTHHNQQQPIILPNLQELILWNMDNMSHVWKCSNWNKFFTLPKQQSESPFHNLTITIKIMYCKSIKY
LFSPLMAELLSNLKHIKIRECDGIGEVVSNRDEDEEMTFTSTHTTTTLFPSLDSLTSFLENLKCIGGG
GAKDEGSNEISFNNTTATTAVLDQFEVCFVHIQLFI.

SEQ ID NO:22

RLG 2B

SEQ ID NO: 23

1 AGTTTTTTTT TTTCCCAATA TCCATTTATA TCGATTTAT TTCTGAAATA ATTTTATCAA AACGCAGGAA
71 ACAATGTAGA ATAATACTGG TATAATTAAT TATATAAAGT TATTAGGCTG AAATCTTGAG GCTACTATAA
141 TTTAATTATC ATAATTTGAA AATCATCAAA TTGTATTCCA TGTATATTTA TGTATCAGA TAATTAATPA
211 TATGTGAGCC ACACAAATCC ACATCATCAG ACACCCACC TTATTGTCCG CTACCTCACC ACTTGCATGA
281 TCCCGACATC TTCCCAACCC CACCGACGAC TTGGGGTCTC CTTAATATAT CAATTATTTT CTGTAAAGTAT
351 TTATTGTGT AAATGTGTAA TGTCATTTTA CCTTTTTTCT AATATATACA GAAACATAAA TTTTAAATGA
421 AATTCAACTG CGTTTCATTC TTGCATTAAA AAAAAAGACT GTACTGTTGT CAATATTTTA CTTATAACCT
491 GATTAAATTAA TTAAAGCGTA ATTGCATAAT TTGCATTAGG TTGTAATTTT GTGTTTTATA GGGAGGGTGA
561 GGGTCACCGG GAATCAAAGC ACTTATGTAA AAGCAGGGGA AATACAAAAA ATTTACTCGA AACAAATTTT
631 ATTCAATTTA AGTGAGATAA TAATGTTCTG ATTAGATTAT GAGAACTAGG AGATTTAAGT GATATATCCC
701 ATTTAAAAGA AATTGCATTA TTAATTTTGG ATCTCTTGAT GATGACAAA TTAACCTCGT ACAGGTTATA
771 TATCATATAC AAAATGAGT GCTATGCTTT CGTTTCCAA AAAGCAATTA TAGTTTACT ACACCTACAA
841 ATTTTAAAG GGGTTAAACA TATCAAAAT CTTCATAAGT AATTATATAA ATATGCATTT AACCTCTAA
911 AGAAAATGCT ACTAAGCTTG GACCATCTCA GAATTACAAT CATACCCTTC CCTCAAAAA AGATTCTGAT
981 ATATCATGTC ATTTGGCATT CATTTCCTTT TCACAATTCA TAGTTCTATT CTCAAAAAAT TCGAGTTCTC
1051 GTATTTGTAA GGAAGATCAG AAGAGACTGT TCACACAGGT ACTCTCTTTT ATTTATTTAT TCACATTCT
1121 ATATGTTTAT GTTTCTTGC TTAATGGTIT CGTCAGTCTA ACTGCGCTTG CTGATTTTAA TTTCTTCACT
1191 TTCTTCCACG GATTTTTTAA ATATTAGTTT TGTGAATGAA CAATTGGTGA AGGAAAGAAA CTTCGGAGTC
1261 TTTTCTAAAG TAAACCTAGA TACTTAGGTT ATAAGGGTAT ATGCTAAAAT GAACTATGCC CATTACCTT
1331 TGCCCTTTCT TTTACTTTTT AGTTTTTAGA ATCCAAGTTT TCATATGTAT CTCGATGTGT GAGAAGAATA
1401 GGCATTAGAA AGGTAAAGGA CGTACATAAA ATTGATTAAAT TAGTGAATGT TCTTTGATAT CATTATTTTT
1471 ACTCTCATAA AAAGCATATA GATCAAACAC AAATTGCTAC TTGTTAGTGT AACAACCTCG ACTTAATAAT
1541 GTTAATATC AAGATTCTCT TGATTTC AAC TATTTTCTAA CCGAACAAGC TCACTAAAAA CTCATATTGC
1611 TTTGAGTCTG AGTGGTTTTAT ATTTGGGGTT TTACATTTAA TTTTGTGTC ATGAATGTGA AAATAGACTG
1681 CTTATTGATT CTTTGTGTTT CATTGAGTTG ATTTTCATTA TTTACTACCTT ACAAAATTCCT CAGTGATAGA
1751 TTTCCAATAA TTTGCTAATT CGTTGCTTTC TAAATATGTA GGAGCTACTA AAAGCAAAAA TATCGAGCAA
1821 TGTCCGCCCC AACCGGGATT GCTGGTGCCA TTTATTAACCC AATTGCTCAG ACGGCCCTTG TTCCCGTTAC
1891 GGACCATGTA GGCTACATGA TTTCTGTCAG AAAATATGTG AGGGTCATGC AGATGAAAAT GACAGAGTTG
1961 AATACCTCAA GAATCAGTGT AGAGGAACAC ATTAGCCGGA ACACAAGAAA TCATCTTCAG TTCCATCTCA
2031 AACTAAGSAA TGGTTGGACC AAGTAGAAGG GATCAGAGCA AATGTGGAAA ACTTTCCGAT TGATGTCATC
2101 ACTTGTGTA GTCTCAGGAT CAGGCACAAG CTTGGACAGA AAGCMTTCAA GATAACTGAG CAGATTGAAA
2171 GTCTAACGAG ACAACTCTCC CTGATCAGTT GGACTGATGA TCCAGTTCYT CTAGGAAGAG TTGGTTCCAT
2241 GAATGCTATC ACCTCTGCAT CATTAAGTGA TGATTTCCCA TCAAGAGAGA AAACCTTTTAC ACAAGCACTA
2311 ATAGCACTCG AACCCAACCA AAAATTCCAC ATGGTAGCCT TGTGTGGGAT GGGTGGAGTG GGGAGACTA
2381 GAATGAAGCA AAGGCTGAAG AAGAAAAGAA ATTGTTAAT TATATTGTTG GGGCAGTTAT
2451 AKGGGAAAAG ACGGACCCCT TTGCCATTCA AGAAGCTATA GCAGATTACC TCGGTATACA ACTCAATGAA
2521 AAAACTAAGC CAGCAAGAGC TGATAAGCTT CGTGAATGGT TCAAAAAGAA TTCAGATGGA GGTAAGACTA
2591 AGTTCCCTCAT AGTACTTGAC GATGTTTGCC AATTAGTTGA TCTTGAAGAT ATTTGGGTTAA GTCCTTTTCC
2661 AAATCAAGGT GTCGACTTCA AGGTCTTGTT GACATCAGCA GACTCACAAG TTTGCACTAT GATGGGGTT
2731 GAAGCTTAAT CAATTATTAA CGTGGGCTTT CTAATCAGAG CAGAAGCTCA AAGTCTGTTT CAACAATTTG
2801 TAGAACTTTC TGAGCCCGAG CTCCAGAAGA TAGGAGAGGA TATCGTAAGG AAGTGTGCG GTCTACCTAT
2871 TGCCATAAAA ACCATGGCAT GTWCTCTTAG AAATAAAGA AAGGATGCAT GGAAGGATGC ACTTTCCGCG
2941 ATAGAGCACT ATGACATTCA CAATGTTGCG CCCAAAGTCT TTGAAACGAG CTACCACAAT CTCCAAGAAG
3011 AGGAGACTAA ATCCACTTTT TTAATGTGTG GTTTGTTTCC CGAAGACTTC GATATTCCTA CTGAGGAGTT
3081 GATGAGGTAT GGATGGGGCT TGAAGCTATT TGATAGAGTT TATACGATTA GAGAAGCAAG AACGAGCTC
3151 AACACCTGCA TTGAGCGACT GGTGCAGACA AATTTGTTAA TTGAAAGTGA TGATGTTGGG TGTGTCAGA
3221 TGCATGATCT GGTCCGTGCT TTTGTTTTGG GTATGTTTTC TGAAGTGGAG CATGCTTCTA TTGTCAACCA
3291 TGGTAATATG CCTGGGTGGC CTGATGAAAA TGATATGATC GTGCACTCTT GCAAAAGAAAT TTCATTAAAC
3361 TGCAAGGGTA TGATTGAGAT TCCAGTAGAT CTCAGTTTCT TTTATGAAGG AATGGAAAAG CTCCATGTTA TATCATAGA
3431 GAGATAAGTC GCTAAGGTTT CCTCAAGACT TTTATGAAGG ACCTCGATGC TCCACCAACA TTCCGGTGCT TCATCTCACT
3501 TAAAATGAAG TACCCATTGC TTCTTTTGGC ATCTATCGAA ATCTATCGAA TCTGGAAGTG CTGAGCTTTG
3571 GAATGTTTCT TAAAGATGTT TGATTGCTCT TCTATCGGAA TTTAAAGAAG CTAAGGTTAC TTGATCTGAG
3641 CAAATCTTCA CATTGAATGG TTACCTTCCA CAGTCAGAAA TTTAAAGAAG AGTTTTGTCA AACTTGAAGA CTTTTATATT
3711 ATTTTGTGAT GGTCTCCGTA TAGAACGGG TGCTTTGAAA AGTTTTGTCA AACTTGAAGA CTTTTATATT
3781 GGAGATGCAT CTGGGTTTAT AGATGATAAC TGCAATGAGA TGGCAGAGCG TTCTTACAAC ATTTTCTCAT
3851 TAGAATTCGC GTTCTTTAAT AACAAGGCTG AAGTGAAAAA TATGTCATTT GAGAATCTTG AACGATTCAA
3921 GATCTCAGTG GGATGCTCTT TTGATGAAAA TATCAATATG AGTAGCCACT CATACGAAAA CATGTTGCAA
3991 TTGGTGACCA ACAAAGGTGA TGTATTAGAC TCTAACTTFA ATGGGTTATT TTTGAAAACA GAGGTGCTTT
4061 TTTTAAAGTG GCATGGCATG AATGATCTTG AAGATGTTGA GGTGAAGTCG ACACATCCTA CTCAGTCTCT

RLG 2B cont.

SEQ ID NO: 23

```

4131 TTCATTCTGC AATTTAAAAG TTCTTATTAT TTCAAAGTGT GTAGAGTTGA GATACCTTTT CAAACTCAAT
4201 CTTCGAAACA CTTTGTCAAG ACTTGAGCAT CTAGAAGTTT GTGAATGTGA GAATATGGAA GAACTCATAC
4271 ATACTGGAAT TGGGGGTTGT GGAGAAGAGA CAATTACTTT CCCTAAGCTG AAGTTTTTAT CTTTGAGTCA
4341 ACTACCGAAG TTATCAAGTT TGTGCCATAA TGTCAACATA ATTGGGCTAC CACATCTCGT AGACTTGATA
4411 CTTAAGGGCA TTCCAGGTTT CACAGTCATT TATCCGCAGA ACAAGTTGCG AACATCTAGT TGTGTAAGG
4481 AAGGGGTAGA TATATGTTCT TTATGTTAAT ACAATTTAAA TAATATTTTC AACCAAATTT TCATAATATA
4551 TCTGTAATTT GATTGTATGA TGTGTTATTG TTTATATGTG GCTATTAGG GATGATTATT TTGCAGGTTG
4621 TGATTCCCTAA GTTGGAGACA CTTCAAATTC ATGACATGGA GAACTTAGAA GAAATATGGC CTGTGAACT
4691 TAGTTCGAGGT GAGAAAAGTTA AGTTGAGAGC GATTAAAGTG AGTAGCTGTG ATAAGCTGTG GAATCTATT
4761 CCGCGCAATC CCATGCTCTCT GTTGTCATCAT CTTGAAGAGC TTACAGTCCA GAATTCGGGT TCCATTGAGT
4831 CGTTATTCAA CATTGACTTG GATTGTGTCTG GTCCAATTGG AGAAGAAGAC AACAGAGGCC TCTTAAGAAG
4901 CATCAACGTG GAGAATTTAG GGAAGCTAAG AGAGGTGTGG AGGATAAAAG GTGCAGATAA CTCTGATCTC
4971 TAGGTATCAA TCAAGCTGTG TGAAAGCATA AAGATTGAAA AATGTAAGAG GTTTAGAAAT ATATTCACAC
5041 CTATCACCGC CAATTTTTAT CTGGAGGCAC TTTTGGAGAT TCAGATAGAA GGTTCGGGAG GAAATCACGA
5111 ATCAGAAGAG CAGGTAACGC TTTCAATTTT ACTTTCTTAA TTAATTAAGG ACTAAGCTCC TGTTTTTTGA
5181 ATAATAAAGA GGTGGGATGA CTAAACTTGG GCATCACAAT TGCAACAAA TGTTACAAAC CATGAAACGT
5251 TCAAACCATT TCTTGAATTA AGGTTTCAAT ACAAGTCATT TAAAAATATG GCTTAAATTT TTTTATATT
5321 TAGGTATCAA CATGATTTTT CATTAGAGAT CATTATTATA ATAGTAAGTT TAAAGCAATT TAAATCAGAA
5391 CTAATTTCTA CTTTAGCTAA TAAATCGTTA TAAATGTAAA TAATTACTTT TTAGTGAAAT AAGCAACGGA
5461 TTTAATAAGT TAACAACCTA AATGTCATTT CCTAACAAAA AAAACTTTTG TTCAGAAAA CCGCAATTCA
5531 AGATAACTAA AATAAAAAATA TTTGACATTC ACTAAGAGCA TTTTTTTTTT TAAATATGAT TGCAAATGAA
5601 TAAAACCTAA ATTTATACAG AAAATCTTTT TATATATGTT ATACAAAATT TACAAATTGA AATTGGATAT
5671 GTTAATTAAAC GGTTTATAAT TCTGGTATCA CAAAGGGATA TATAATAAAA TATTATTTTC TGTAGTCAAT
5741 TGTAATTGTA CTAGTTTATA ACCCGTGGGA ACCATGAGTT CTAAAATTAG TTAACCTTTC ATAAATAAAA
5811 TTTATAATTA TTATTTATTT TAAATAAATT ATTAATTAAG AGATATATCA AAAATTTAAA GTTATTATAA
5881 CTTCAAATTT AACATATAAT TAGAAAAATC ATGATCATAA CTCTGCACT CTCTTTGTAT AAATGCAGAG
5951 AAGCTATTAG TATATTTCCTA ATCAAGTCCA AGCCTAATGA AGCCTATATA ATTTTGTGAA AACTCAATTA
6021 GCATTAGGTT TTAAGAGTCA CCAAAITCAA AGAATAATCC AATGCTTTCA TTACCACTAT GGAGAAAATA
6091 TTTCTTAGT TTAATGAAA TGAAAACAAA CATTCAAAC AATTGTTGCT TATTAAACCA AAGACCCATT
6161 ACTTAGCCAA GAGTTTAACA AAAAAAATT ACATTCATGT ATCATTATTC ATGACTAGAT ATATATGAAC
6231 ATGAAGGAG TTTTATAGA AAATATAATC ATAGATATTC AACATAACTT CAGGGAATTC CTCAAAAATA
6301 CCAAGTTATT CAGAAAATTA CATCCAAAGT AACCAGAGAG AAGTTTAGCC TAGCATGGCT AAACCTAAGA
6371 AACTAAAATA AGGATTAGAA GTACCAAACA TGTAAGTAA ATCACAGTAA AAGATGATGT TGTCTTGAT
6441 GTTCTTCTAA GTTCTTCAAG TCTCCAGTTG CTCTTAATAA TGCAAAGGAG AGCCATTAAA TTCGTATGTA
6511 TTGATCCCTT CAAAAGCTGC ACCAACCTCC CTTAAATAAC ACTCAAAGCA AAAATGACAA AATGCCCTGA
6581 AGGACCTAT GTGGGTGCC TGGCGGGTG GCATGCTTCC GCGCGTCTCA CACAGGTGAT CACAGGTGAT
6651 GTTGTGCGGG ATAGCTTGTC GCATGCTTCC GCGCGTCTCA GAGTCCAATC AGCTTATAAC
6721 TGGCTTCTTG AGTTTGTAGC CTCCGATGCT TAGTCCACTT GGCCCAATTC GAGTCCAATC AGCTTATAAC
6791 CCATTTTTCT TCAAGTTATC TTCAAGTTAA GCCCAATTTG GCTTCTCCAA ATCATCCATA ACTTCACAGA
6861 ATCGCCCGTT CATCTTAATC CCGGATGCAC AATPATCTCT CCGTCTTCAT TTTAAGCAAG ATACCACCTT
6931 CTTCAATGCT CATCCATCAA TAGTACACTT CATGTATCAT CTCTACTAGT TATTAGTCC ACAAACTCTT
7001 GTTGTCTCTC AAATTTAATT ATCTCATTTA GTTCCCCGTT CCGTACTTT CTTTAAATTT TGGAAATTAAG
7071 CTCAGAGAAA TATTAAGTAC CCGAAATGGT CATAAAATTA ACAAAGGGA AAATGCATGA AGATTAACTA
7141 AATGATGAAC GAAATATGCT AAAATAGACT ATAAAATGAA GTAAAATAAA TGAAATATAT GCACCTCCGAC
7211 CACCTTTATG GCTTGTAGTC CACCCACCTT TCATTCTTTG TACCAATATG GGATGGAAC ATCATTAATT
7281 AAGCCAAAAA GCTAACATAT AAGGGTTTAG TGACAAAAGG ATGACAAAATA ATCCATTTTT
7351 CTTGTTTTTA CACAACACAC ACATAGGGGC AGACGTAGGA TTTCAAAGTA CAGATTGTGT GTGGCACATA
7421 AGTGTGCTG GTGACATTTT TTTTTCTTTT TTACGTGGTG GCACAACAGT AGGAAAAACG AAAAATTCGA
7491 AATTTTTTAC AATTTGTCTT AAAAAAACA GGGGTGTGTG GTGCCACTAT GGACAACAAA GTTGAACCTG
7561 CTTACCGCGG CACACACACA GAGAGAGAGA GAGAGAGAGA GAGAGAGAGA AAGAAAGAAA
7631 GAGAGAGAGA GTTTGGGATG TGATACTTCT TTTAGGAAAA TGGAGTTATA TCTTTGATAT TGTATTTTTT
7701 TAATGTAATT TATNTATTTA ATCATTTTAG TTTATAAGTT NTATTTTATN GGNTATGAAA AAAAAAGTCT
7771 TTTATACATT GGATTTAACA TAAAAATCCA ACAATATTAA TCAAAAGAC CAAACATGTG GACAATTATG
7841 TATATAATTA ATTCACAATA GTCTTTAGGA ATAGTATTAT ATATATAATT AATTCTCAAT GGTCTTAGGA
7911 ATAGTTAAGT CTATATTTTC AAACCTTTTC CACAATCTTT TGCTTACTTT GACACTTTTC TCTATTATTC
7981 TTACATATAT ATATATATTA AAGCGCAAAG GTCATAGGAA TATAATATTT TATTAATTTT TCTATTATTC
8051 CACAAAAGTT TGAACACTTT GCCACTTTTT GTCCCTCTCT AACCTTTTCA ATGTTTTGCG ACAAAAAGTTC
8121 CAAAACCTTG CCACTTTGAT CATTCCTCAA CTTTTCACCG CATTAGTTTG TGGAGTTGGC AGTTTTGGTC
8191 CCTCTAACTT CGATATTCTC TACTGCTAGC CAAAAGGGT TCCAGAGTTT CACACTTTTG GTCCCTGACA

```

RLG 2B cont.

8261 GTAACCAAAT GTGAGATGTC AAATTTTTC CACATTAGTT TGTGGAGTTG TCCCTTTTGG TCCCCCACA
8331 TTCGATATTC TACTATACGA TCTTATTTT CTCAAATAAC AACACGTATA TTTCATC:CT AATTGAAAA
8401 AGAGTTTTAA AA:AAATAAC GACTAGG::: G:GC:GAGTT TTTTTT:ACA AGTTTGTATC AAATCATATC
8471 AAAATTTAAG GTGGAACGGT GACCACATTA ACCAGAAATG TAATTTATTC TTTGATTTTG ATAATTTTTA
8541 ATATTTTGTG GTGATCTATG TATTTAAAAG TAAACAACAA AGAACATAAT CCAAAAACCT AAATGCAAG
8611 TCTCGCCCAA TTTCTCTATC ACTAGTCCTC ACTTACGATG GCGTTACGTC GCTCTCTCAC TGCTTACAAC
8681 CCTTTGTGTC TACTCATTTAC AATAACGAAA AGTTGAATAT CCATATATTT ATTTGGATGT GGAATTGAAC
8751 GAATCTCGTC AAAATTTTGA TTTTGTGAT GGATTGAGT AGAAGTTTGG GCAGAACGGG AATGATGGTC
8821 TGCAAGTGGT TATAAACTTG ATTCTGAGTT ATTACTATAT ATGTAGCCTC TTTACAACGA CCAAGGTTTC
8891 TTCCAGGTAC CATTGATCT TTTTAGAACT TAGTTTTCTG AAACACCTCG ATTTGGATCA AATATCACCA
8961 ACAACTCTTA AAAACTTGAT TAATCAATTG TTTTCTTCAT CTTGATAACA AGTGGAATGA TTTTCTACTT
9031 AGATTAACTT GAAAAAAAG GTCCATGTGC GTCTGGTGA TCTGGTAAAT GAAGATGGAA GGGAGAGCTG
9101 ACTTTAAAGA CACAAACACG TCACCATATC TCTTATTTTA TTTTAAATTT GCTTTTGGTG TATTTTCTTT
9171 TTCCCTATTT CTTTCTTTCT TGATCTCCAG ATGGTATG TGTTGGATAA TTTACACCTA GAGATTGGGA
9241 ACGATGGGAA GGGGTCTGTG ATTTATGGCT GGCCGAGTTT TACTTATTAA ACCTAAATTC
9311 TGATCTCTGT TTGAAAATAA GTTGCACTTT TATTTTTGTA TTATCTGTG GCATAGGATC CTTAGCATCT
9381 TTTAATAATT TATTTGAAGG TGAAAGATCC AACTATTTTT TAGCTGTGG CATTTTCCAT CATTTGCAAC
9451 TGTTCTTTGA AAAAAAATA CCTAAAATAA AAATAACCAT TTTCAAATCC AAAATTATAA GAGAGAATTG
9521 TAAATGGACA TGGAAATCATA AATCATTAAC ACAGTTCAGT AAACAAGTTG CTAATTACAT TTTGTCTGT
9591 GCAGATTGAA ATTCTATCAG AGAAAGAGAC ATTACAAGAA GCCACTGGCA GTATTTCAAA TCTGTATTC
9661 CCATCCTGTC TCATGCACCT TTTTCATAAC CTCGTGTGTC TTACATTGGA TAATTATGAA GGAGTGGAGG
9731 TGGTATTTGA GATAGAGAGT GAGAGTCCAA CATGTAGAGA ATTTGGTAACA ACTCGCAATA ACCAACAACA
9801 GCCTATTATA CTTCCCTACC TCCAGGATTT GTATCTAAGG AATATGGACA ACACGAGTCA TGTGTGGAAG
9871 TGCAGCAACT GGAATAAATT CTTCACTCTT CCAAAACAAC AATCAGAATC CCCATTCCAC AACCTCACAA
9941 CCATAAATAT TCTTAAATGC AAAAGCATT AAGTACTGTT TTCGCCTCTC ATGGCAGAAC TTCCTTCCAA
10011 CCTAAAGGAT ATCCGGATAA GTGAGTGTGA TGGTATTAAA GAAGTTGTTT CAAACAGAGA TGATGAGGAT
10081 GAAGAAATGA CTACATTTAC ATCTACCCAC ACAACCACCA CTTTGTTCCT TAGCTTTGAT TCTCTCACTC
10151 TAAGTTTCCT GGAGAATCTG AAGTGTATTG GTGGAAGTGG TGCCAAGGAT GAGGGGAGCA ATGAAATATC
10221 TTTCAATAAT ACCACTGCAA CTACTGCTGT TCTTGATCAA TTTGAAGTAT GCTTTGTACA TATTCCATTA
10291 TTTATTTAAT TTCTTTTTTT ATTTGCAATA TTCTATAAAT AATACATTTT ATACCCACTA TACTAAGATA
10361 ATAATTACCT AGAGGGATGG ATGCTATGAC ACAGCTGCTA CACTTCAGAA ACTCTARTAA GGGCAGTTAT
10431 GGAAGTTCAA TAAAAATGATA ATGGCATCTT TTGATGGGTA ATATAGGCAA TTTAAGTTTT ATTTCTGTTA
10501 AAGCAGTATT TAGCAAGTAC TGGCCAGTAG GAGAGGAGAA TATCACCTTT TGIGAAAATC TGGTCATTGT
10571 ACCCAGAATT TAGTTAAATG TAACATTTTA GATATTAGGG GTTATCAGGT GACAGATATT GTAGAATAGA
10641 ACAATATGTA ATATTACCCA AAACATTTTT TTCTAAGGTT GCTCTGTAA ATATGTGCTT TCTTGATTTT
10711 ATTGAATTTG CATTCCTATA TTTTAGGTGG TAAAGTGATT GTCTCTTCAA TAAATCCCGA AATTTTTTAA
10781 TTAACAAAAA AAAAAACAAA AGTAAATTTT TGATATGGAG AGCACTGGTA TCATTTAGTA TATAAAAAAC
10851 AGATTTTGAA TTAAGTTTCT TATATAAAG CTGTGTATAT AGTTTAATTA GTTTTACATC ATTTTCCAT
10921 GTGGTGTGTC AGTTGTCTGA AGCAGGTGGT GTTCTTGGG GCTTATGCCA ATACGCTAGA GAGATAAAAA
10991 TAGGCAACTG CCATGCATTG TCAAGTGTGA TTCCATGTTA TGCAGCAGTA CAAATGCAGA AAGCTT

SEQ ID NO: 23

RLG2B a.u.

MSDPTGIAGAIINPIAQTALVPVTDHVGYMISCRKYVRVMQMKMTELNTSRISVEEHISRNTRNHLQIP
SQTKEWLDQVEGIRANVENFPIDVITCCSLRIRHKLQKAFKITEQIESLTRQLSLISWTDDPV?LG RVG
SMNASTSASLSDDFPSREKTFTQALIALEPNQKFHMFVALCGMGGVGKTRMMQRLKKA?EEKLFFNYIV
GAVI?EKTDPFAIQEAIADYLG IQLNEKTKPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGL
SPFPNQGVDFKVLTSRDSQVCTMMGVEANSIINVGLL TEAEAQSLFQQFVETSEPELQKIGEDIVRKC
CGLPIAIKTMAC?LRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQEEETKSTFLMCGLPEDFDI
PTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQTNLLIESDDVGCVKMHDLVRAFVLGMFSEVEH
ASIVNHGNMPPGWPDENDMIVHSCKRISLTCKGMIEIPVDLKFPKLTILKLMHGDKSLRFPQDFYEGMEKL
HVISYDKMKYPLLPLAPRCSTNIRVLHLTECSLKMFDCCSIGNLSNLEVL SFANSHIEWLPSTVRNLKKL
RLDLRFCDGLRIEQGVLSFVKLEEFYIGDASGFDDNCCNEMAERSYNLSALEFAFFNKA EVKNMSFE
NLERFKISVGCSDENINMSSH SYENMLQLVTNKG DVLD SKLNGLFLKTEVLFLSVHGMNDLEDVEVKS
THPTQSSSFCNLKVLISKCVELRYLFKLN LANTLSRLEHLEVCECENMEELIHTGIGGCCGEETITFPKLF
LSLSQLPKLSSLCHNVNIIGLPHLVDLILKGIPGFTVIYPQNKLRTSSLLKEGVVIPKLET LQIDDMENLEE
IWPCELSGGEKVKLRAIKVSSCDKLVNLFPRNPMSLLHHLEELTVENCGSIESLFNIDLDCVGAIGEEDN
KSLLRSINVENLGKLREVWRIKGADNSDLINGFQAVESIKIEKCKRFRNIFTPTANFYLEALLEIQIEGCG
GNHESEEQVTL SISLS

SEQ ID NO: 24

→ 25

[illegible]

RLG2A
RLG2B
RLG2C
RLG2D
RLG2E
RLG2F
RLG2G
RLG2H
RLG2I
RLG2J
RLG2K
RLG2L
RLG2M

410 420 430 440 450 460 470 480 490 500
 AGCAGAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG
 .AACAGAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TTTATCTTACCTTCCCA-TAATATAGGAGTATATTTGTAAAGGAG 487
 AGCAGAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-GAAGATAGGAGAGATATTTGTAAAGGAG 485
 AGCAGAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-CAAGATAGGAGAGATATTTGTAAAGGAG 477
 ANCANRANCAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 455
 AGCAGAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-CAAGATAGGAGAGATATTTGTAAAGGAG 480
 AGCTGAAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 485
 TGTGAAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 489
 TGTGAAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 467
 TGTGAAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 481
 TGTGAAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 464
 AGAAGAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 484
 AGTGAAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 430
 AGCAGAGGACAAAGTTTGTTCACCAATTTGTAGAACTTC-----TGAGCCCGAGCTCCA-TAAGATAGGAGAGATATTTGTAAAGGAG 459

RLG2A
 RLG2B
 RLG2C
 RLG2D
 RLG2E
 RLG2F
 RLG2G
 RLG2H
 RLG2I
 RLG2J
 RLG2K
 RLG2L
 RLG2M

510 520 530 540 550 560 570 580 590 600
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 586
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 584
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 576
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 554
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 579
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 584
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 588
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 566
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 580
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 564
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 583
 TGTTCG-CGGTCTACCCATTCGCATCAAAACCATGGCTGTACTCTTAGAATAAAGAGAGATGCAATGCACTTTCTCGTTTGTAGAGCAACCAT 529

RLG2A
 RLG2B
 RLG2C
 RLG2D
 RLG2E
 RLG2F
 RLG2G
 RLG2H
 RLG2I
 RLG2J
 RLG2K
 RLG2L
 RLG2M

RLG2A
RLG2B
RLG2C
RLG2D
RLG2E
RLG2F
RLG2G
RLG2H
RLG2I
RLG2J
RLG2K
RLG2L
RLG2M

RLG2A
RLG2B
RLG2C
RLG2D
RLG2E
RLG2F
RLG2G
RLG2H
RLG2I
RLG2J
RLG2K
RLG2L
RLG2M

1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1101
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT
AAGACTTTTCATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1071
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1075
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1067
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1045
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1064
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1069
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1082
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1080
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1074
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1058
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1077
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1014
AAGAAATTTCAATTAACATCCCAAGGGTATGCTGAGTTTCCAGTAGACCTCAAGTTTCCAACTTAACGATTTTCAAACTTTGATGCAATGAGATTAAGTCGCT 1049

RLG2A
RLG2B
RLG2C
RLG2D
RLG2E
RLG2F
RLG2G
RLG2H
RLG2I
RLG2J
RLG2K
RLG2L
RLG2M

1110 1120 1130 1140 1150 1160 1170 1180 1190 1200 1201
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC
GAGGTTTCCCAAAACCTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1171
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1175
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1167
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1145
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1164
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1169
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1182
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1160
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1174
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1158
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1177
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1114
AAGGTTTCCCTCAAGACTTTTATGAAGGAATGGAAAGGCTTCAGGTTTATATCATACGATTAAGTATCCATTCCTTCCCTGCTCACTCAATGCTCC 1149

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
 CTTAAATAATTTGGTCAAACTTGGAGGAGCTCTATAT-GACAGTG-----GTT-----CGAGGTGGAAGG-----CGA----- 1437
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATCTGGTTTATAGATGATTAACCTGAAATGAGATGCGAGAGCGTTCTTACACCTT 1475
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1417
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1411
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1412
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1413
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1414
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1415
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1416
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1417
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1418
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1419
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1420
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1421
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1422
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1423
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1424
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1425
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1426
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1427
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1428
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1429
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1430
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1431
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1432
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1433
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1434
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1435
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1436
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1437
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1438
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1439
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1440
 CTTGAAAAGCTTTTGTCAAACTTGAAGATTTTATATATGGAGATGCAATG 1441

XXXXXXXXXXXXXXXXXXXXXXXX
 1510 1520

1439 1500 1417 1412 1432 1438 1466 1420 1442 1474 1437 1398 1441
 TCTGCATTTAGAAATTCGGTCTCTTTA - TTA - SEQ ID NO: 27
 - SEQ ID NO: 28
 - SEQ ID NO: 29
 - SEQ ID NO: 30
 - SEQ ID NO: 31
 - SEQ ID NO: 32
 - SEQ ID NO: 33
 - SEQ ID NO: 34
 - SEQ ID NO: 35
 TGGAAATGGCGAGCCTTTTTCACA - SEQ ID NO: 36
 - SEQ ID NO: 37
 - SEQ ID NO: 38
 - SEQ ID NO: 39

	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919
--	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

SEQ ID NO:

[illegible]

	110	120	130	140	150	160	170	180	190	200
AC15-2A	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	
AC15-2B	ATCTCTGTTTGA	AAAGTAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	187
AC15-2C	ATCTCTGTTTGA	AAAGTAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	187
AC15-2D	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	186
AC15-2E	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	188
AC15-2F	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	187
AC15-2G	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	187
AC15-2H	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	176
AC15-2I	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	164
AC15-2J	CCCATCTTGATTA	CCACCAATG	GAAGCAACGAT	CAAAACAGGCGCT	TATACA--	-ATTTG--	TTGTGGTTTGTATTTATTT	TCGGTGAAGCTG		194
AC15-2L	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	122
AC15-2N	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	187
AC15-2O	ATCTCTGTTTGA	AAATAAGTT	-GCATCTTTATTT	-TG--	TATTATCTTGTTGCATAGGATCCT	-TAGC	ATCTTTTAATAATTTATTT	-----	GAAGCTG	168

AC15-2A	210	220	230	240	250	260	270	280	290	300
AC15-2B	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	286								
AC15-2C	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	284								
AC15-2D	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	285								
AC15-2E	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	286								
AC15-2G	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	287								
AC15-2H	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	288								
AC15-2I	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	289								
AC15-2J	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	290								
AC15-2L	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	291								
AC15-2N	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	292								
AC15-2O	AAAGATCCAACTATTTTAACTGTTGGCAATTTCCATCTTTGCACTGTTTCTGCAAAA--TACCTTAAATCAAAATTAACCATTTTCAAAATCCA	293								
AC15-2A	310	320	330	340	350	360	370	380	390	400
AC15-2B	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	375								
AC15-2C	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	373								
AC15-2D	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	374								
AC15-2E	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	375								
AC15-2G	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	376								
AC15-2H	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	377								
AC15-2I	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	378								
AC15-2J	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	379								
AC15-2L	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	380								
AC15-2N	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	381								
AC15-2O	AAATTTAAGAGAGAAATTTGTTAAATGGAC--ATGGAAATCAATAAATCAATTAAC---ACAGTTCTAGTACACAGTTGCTTAATTAACATTTCTTCTGCTG	382								

SEQ ID NO:

AC15-2A
AC15-2B
AC15-2C
AC15-2D
AC15-2E
AC15-2G
AC15-2H
AC15-2I
AC15-2J
AC15-2L
AC15-2N
AC15-2O

810 820
TAAGTACTTGTGTTTTCACCTCTCAGG -56
TAAGTACTTGTGTTTTCACCTCTCAGG -57
TAGGTACTTGTGTTTTCACCTCTCAGG -58
TAGGTACTTGTGTTTTCACCTCTCAGG -59
TAAGTACTTGTGTTTTCACCTCTCAGG -60
TAAGTACTTGTGTTTTCACCTCTCAGG -61
TAGGTACTTGTGTTTTCACCTCTCAGG -62
TAAGTACTTGTGTTTTCACCTCTCAGG -63
TAAGTACTTGTGTTTTCACCTCTCAGG -64
TAGGTACTTGTGTTTTCACCTCTCAGG -65
TAAGTACTTGTGTTTTCACCTCTCAGG -66
TAGGTACTTGTGTTTTCACCTCTCAGG -67

779
777
777
788
721
781
738
722
784
699
778
763

SEQ ID NO: 68

RLG3 (real RLG3)

[Strand]

```
1  AATGGCAAAA GAAGTCGGAG CAAGAGCTAA GTTAGAGCAT CTATTTGACG TCATTATCAT GGTAGATGTC
71  ACTCAAGCAC CCAACAAGAA CACAATTCAA AGTAGTATTT CAGAACAGTT GGGATTAAAA CTGCAAGAAG
141 AGAGCTTGTT GGTAAAGAGCA GCTAGGGTAA GTGCGAGGTT AAAAATGCTT ACAAGGGTGC TGGTGATATT
211 AGACGATATA TGGTCAAGGC TTGACATGGA GGAAC TTGGG ATTCCCTTTG GATCAGATAG ACAACACCAC
281 GGCTGCAAAA TCTTGTGAC TTCAAGAAGT ATTAGTGCTT GTAACCAGAT GAGAGCTGAT AGAATCTTTA
351 AAATACGAGA AATGCCACTG AATGAAGCAT GGCTTC TTTT CGAAAGAACA GCTAAAAAAG CTCCGAATCT
421 GCATCAAGTA GCAAGAGATA TCGTGGAGGA GTGTGGTGGG C
```

RLG4
SEQ ID NO: 69

```
1  GAATTCGGTG  TTGGTAAGAC  AACTCTTGCC  TCTTCTGTTT  ATGATGAAAT  CTCTAGCAAG  TTTGATGGTT
71  GCTGCTTTCT  AAAAATATCT  GGGAGGAATC  AAGTAATAAA  GACGGTATAG  AAAGATTGCA  AGAAAAAATC
141  ATTTGTGATG  TTTTGAAACA  AGAGCAAGTG  GGGTAGGGA  GAGTTGAAGA  AGGAAAGCCG  ATGATAAAGG
211  ATAGGTTACA  ACATAGAAAG  GTATTGATTG  TGCTTGATGA  TGTCGACAAC  GTTGAGCAGC  TAGCTAGAAC
281  AGTTGGCTGG  ATCACATGAT  TGGTTTGGTG  AAGGTAGCCG  CATAATAATC  ACAACTAGAG  ATGAACATGT
351  ATTAATTGCA  CACAAAGTAG  ATGTGATACA  CAATATAAGC  TTGTTAAACA  ACGATGAAGC  TATGCATCTC
421  TTCTGCAAGC  AAGCACCACG  GGGTCACAAA  CGTATACAAG  ATTATGAGCA  ACTTTTAAAA  CATGTGGTTT
491  CTTATGCTGG  TGGGCTTCCA  CTAGCACTGT  CGAC
```

SEQ. ID NO: 70

RLG1-E169

[Strand]

```
1  ATCGTAACCG  TTCGTACGAG  ANCGCTGTCC  CTCCTTCATC  TTTTGTGATA  TGTCATATTC  TCATNATTN
71  TGCCACATAT  AATTTTGTGG  TTATTTTAAA  TTAATTTTAA  TTCCACATGT  CATTTTATGA  GTTTTCTAT
141  TTTATTGAGT  TTCACATAAT  ATTTAAATGT  AATAACAATA  AATGCATATT  TATTTTCTTT  TAAATAAACG
211  CATATAATAT  ATAGATTAAA  ATCATATAAT  ACATAGGTTA  AACTCATATA  ATACATATGT  TCATCCCCAG
281  TTTATTTTAA  TGTCTCATCC  TTAATTTATT  TATTATTATT  TTATTAGAGT  AGATGATCTT  TGTGATATTA
351  AAAATTTAAT  TTGTTCAAAA  TTTAAAATTA  TTAATAATCC  CACAATTTGA  ATAAAATTAA  AAAAAATGGN
421  CCCACCATTA  TTTCCATCACT  TTTTCAGCTC  ATCAATATCG  TGAGTATTCT  CCTTCGTTTC  CACCCTAATC
491  AATATTTCCA  GCGAATGACA  GACTCCCTACG  GCGTTTCTGA  ATTTGCGTTC  CGACACTGTT  CATTTGAAGGA
561  GATAATAAAT  CAAATGGAGC  TGCTCCAATG  TTCATTGCTG  ATGAAAGGTG  AATTGTATGT  GAAGANAATG
631  TCAGCGATCN  ATCTCCATCC  GGAACCCACC  ACATTATCAG  TGTACCACCA  AACCCTCAA  AACGGYGGAA
701  GTAGRRKAC  WKAAAGTCA  TGAAGAATAG  ATTATTTTTC  TCCTCATGGG  CTGACTGAGG  AGCGGGTTTA
771  GTTCATCAAT  TTTCTTTGAN  CAAAGAATTA  TCGGTCCATC  TTTTATCTTT  TTCTGTGAAA  CTCCTCAATT
841  TCGCAATGTT  TTGTTAAACA  ATTTTAAATC  TTTTATCTTT  TTTTATCTTT  TTTTATCTTT  TTTTATCTTT
911  ACTTGCAACT  TTTGGGCCCA  CAAATTTGTG  GTGGGCGTTA  ATTTAATCCA  CATATTCAC  GTAAACAATA
981  ATTCAAATCG  ATCTCTGTTT  ATCCAAATCA  TCAACATCTC  TTGATAATTG  AAATCATTTA  CGCTTCATCC
1051  ATTTTATCCA  CATCTATACT  ATATTCTCTG  CTCCTATCAT  ATTAACGAT  GGCTGAAATC  GTTCTTCTG
1121  CCTTCTTGAC  AGTGGTGTTC  GAAAAGCTGG  CATTTGAAGC  CTGGAAGAAG  ATTGTTGCTT  CCAAAAGAAAT
1191  TGAATCTGAG  CTTAAGAAAT  TGAAGGAGAC  ATTAGACCAA  ATCCAAGATC  TGCTTAACGA  TGCTTCCAG
1261  AAGGAAGTAA  CTAATGAAGC  CGTTAAAGA  TGCTGAATG  ATCTCCAACA  TTTGGCTTAT  GACATAGAC
1331  ACCTACTTGA  TGATTTTGCA  ACTGAAGCTG  TTCACGCTGA  GTTGACCGAG  GAGGGTGGAG  CCTCCTCCAG
1401  TATGGTAAGA  AAATAATCC  CAAAGTGTG  CACAAGTTTC  TCACAAAGTA  ATAGGATGCA  TGCCAAAGTTA
1471  GATGATATTG  CCACCAAGTT  ACAAGAAGCT  GTAGAGGCAA  AAAATAATCT  TGGTTAAGT  GTGATAACAT
1541  ATGAAAAGCC  AAAAATTGAA  AGGTATGAGG  CGTCTTTGGT  AGATGAAAGC  GGTACTGTGC  GACGTGAAGA
1611  TGATAAGAAA  AAATTCCTGG  AGAAGCTGTT  GGGGGATAAA  GATGAATCAG  GGAGTCAAAA  CTTACGATC
1681  GTGCCCATAG  TTGGTATGGG  TGGAGTTGGT  AAAACAATCT  TAGCTAGACT  TTTGTATGAT  GAAAAGAAAG
1751  TGAAGGATCA  CTTGCAATC  AGGGCTTGGG  TTTGTGTTTC  TGATGAGTTC  AGTGTTCCTA  ATATAAGCAG
1821  AGTTATTTAT  CAATCTGTGA  CTGGGGAAAA  GAAGGAGTTT  GAAGACTTAA  ATCTGCTTCA  AGAAGCTCTT
1891  AAAGAGAAAC  TTAGGAACCA  GCTATTCTTA  ATAGTTTGG  ATGATGTGTG  GTCTGAAAGC  TATGGTGATT
1961  GGGAGAAAT  AGTGGGCCCA  TTCTTGGCG  GGTCTCCTGG  AAGTAGAATA  ATCATGACAA  CTGGGAAGGA
2031  GCAATTTGCT  AGAAAGCTGG  GCTTTTCTCA  TCAAGACCTT  CTGGAGGGTC  TATCACAAGA  TGATGCTTTG
2101  TCTTTTGTTC  CTCAACACGC  ATTTGGTGTA  CCAAACCTTT  ATTCAATCC  AACACTAAGG  CCACATGGAG
2171  AACTGTTTCT  GAAGAAATGT  GATGGCTTAC  CTCTAGCTTT  AAGAACACTT  GGAAGGTTAT  TAAGGACAAA
2241  AACAGACGAG  GAACAATGGA  AGGAGCTGTT  GGATAGTGAG  ATATGGAGGT  TAGGAAAGAG  CGATGAGATT
2311  GTTCCGGCTC  TTAGACTAAG  CTCAATGAT  CTTTCTGCG  CTTTGAAGCT  RTTTRTTTGA  TAYTGCTCCT
2381  TGTTTCCCAA  GGAATATGAG  TTTGACAAGG  AGGAGTTGAT  TCTATTTGTT  ATGGCAGAAG  GGTTTTGTGA
2451  CCAACCAACT  AYAACAAGT  CAAAGCAAGC  KITGGGTCTT  GAATATTTTR  AAGAGTTRTT  GTCAAGRTCR
2521  TTTTTCACAC  ATGCTCTTAA  TRRCAATCS  TTGTTTGTGA  TGCATGACCT  AATGAATGAT  TTTGGTACAT
2591  TTGTTGCTCG  AGAATTTTTC  TCAAGGTTAG  ACATAGAGAT  GAAGAAGGAA  TTTAGGATGS  AATCTTTGGA
2661  RAAGCACCC  CATATGTCAT  TTGTATGTGA  GRATACATA  GGTTACAAAA  RGTTCGAGCC  ATTTAGAGGA
2731  GCTAAAAAT  TGAGAACATT  TTTAGCATTG  TCTGTTGGGG  TGGTAGAAGA  TTGGAAGATG  TTTTACTTAT
2801  CAAACAAGGT  CTTGAATGAC  WTACTTCARG  ATTTACCATT  GTTAAGGGTC  CTRAKTTTGA  TTRRTCTTAY
2871  AATAASYRAG  GTACCARAAK  TCGTSGGTAG  TATGAASCAC  TTGCGGTATC  TTAATCTATC  WGRAACTTWA
2941  ATCACMCAT  TACCGGAAWA  TRTCTGCAAT  CTTTATAATT  TACARACCTT  GATTGTNTCT  GGCTGTGAMT
3011  ATTTAGTTAA  KTTGCCCAAR  ACCTTCTCAA  ASCTTAAAAA  TTTGCASCAT  TTTGACATGA  GGGRTACTCC
3081  KAATTTTRAR  AACATGCCCT  TARGGATTGG  TGAATTTGAA  ARTCTACAAA  CTCTCTTTC  TAACATTTGGC
3151  ATAGCAATTA  CCGAGCTTAA  GAACTTGCAM  AAYCTCCATG  GGAAATTTTG  TATTGGCGGG  CTGGGAAAAA
3221  TGGAAAAATG  MGTGGATGTC  ACGTTAAGCG  AACTTGTCTC  A: AAAAAAGT  TWAATGARTT  ANAACTGGR
3291  WTKGGGGGTG  ATRAATTTAA  TGTTTTCCGA  AATGGGAACA  CTTGAAAAAA  NAAGTCTCTC  AATGAATTGA
3361  ATGCCCTACA  ATGGTAYTCY  AAMWAARRRY  YYWTARWAT  TWMKAWRRK  GKGTTYATRR  TKTMYRAAW
3431  WAGRGTKTK  KARGTAGGTT  TCATCCAATC  ACCCAAGTGG  GAAAATAGAT  GATATTTTGA  GGCCTACTG
3501  ATGAGATGTC  GAGAGGTATG  ATAGGNTNVC  TTGGGGCGGT  AGAAGAAATA  AGCATCCATT  CTTGTAATGA
3571  AATAAGATA  YTGTTGGAAT  CAGAAGCAGA  GGCAAGTAAG  GTTCTTATGA  ATTTAAAGAA  GTTGGATTTA
3641  GGTGAATGTC  AAAATTTGGT  GAGTTTAGGG  GAGAAAAAGG  AGGATAATCA  TAATATTAAT  AGTGGGAGCA
3711  GCCTAACAT  TTTTAGGAGG  TTGAATGTAT  GGAGATGTAA  CAGCTTGGAG  CATTGCAGGT  GTCCAGATAG
3781  CATGGAGAA  TTGTATATCC  ACATGTGTGA  TTCAATNACA  TCCGTCTCCT  TCCCAACAGG  AGGAGGACAG
3851  AAGATCAAGT  CACTTACCAT  CACTGATTTC  AAGAAGCTTT  CGGAAGAGGA  GTTGGGAGGA  CGAGAGAGGA
3921  CAAGAGTCT  TATAAATCA  AAAATTCAGA  TGCTTGAATC  AGTAGATATA  CGTAATTTGG  CAAATCTGAA
3991  ATCTATCAT  GAATTGAGTT  GCTTCATTCA  CCTGAACAGA  TTATATATAT  CAAACTGTCC  GAGTRTGGAG
4061  TCATTTCTCT  ACCATGAGTT  GCCAAATCTC  ACCTCCTTAA  CAGATCGAAG  GAGAGGACAG  CGATTTTCTG
```

RLG1-E169
[Strand]

4131 ACGAACGGTT ACGATTGAC TGGCGTCGT TTT

SEQ ID NO: 70

Further Characterization of RG2 Family Members:

Further sequencing of cloned RG2 polynucleotide sequences, as discussed above, identified additional RG2 species, listed below. Additionally, further sequencing of the 5' sections of RG2 sequences listed above resulted in modified and/or new sequence information, also listed below. The AC15 sequences found in the 3' sections of RG2 family have not changed.

Listed below are: four full length species, RG2A, RG2B, RG2C and RG2S; two near complete, but with a gap in the largest intron, RG2D and RG2J; three nearly complete RG2 gene sequences, RG2K, RG2N, and RG2O. The deduced translation products (polypeptides) encoded by these RG2 species are listed below. The polynucleotide sequences do not contain any gaps (as with some of the polynucleotide sequences), because all of the gaps in the sequences are in introns, *i.e.*, there are no gaps in exon, or coding, sequences.

They include: an RG2A polynucleotide sequence (SEQ ID NO:87) and its deduced polypeptide sequence (SEQ ID NO:88); an RG2B polynucleotide sequence (SEQ ID NO:89) and its deduced polypeptide sequence (SEQ ID NO:90); an RG2C polynucleotide sequence (SEQ ID NO:91) and its deduced polypeptide sequence (SEQ ID NO:92); an RG2D polynucleotide sequence (SEQ ID NO:93) and (SEQ ID NO:94), and its deduced polypeptide sequence (SEQ ID NO:95); an RG2E polynucleotide sequence (SEQ ID NO:96) and its deduced polypeptide sequence (SEQ ID NO:97); an RG2F polynucleotide sequence (SEQ ID NO:98) and its deduced polypeptide sequence (SEQ ID NO:99); an RG2G polynucleotide sequence (SEQ ID NO:100) and its deduced polypeptide sequence (SEQ ID NO:101); an RG2H polynucleotide sequence (SEQ ID NO:102) and its deduced polypeptide sequence (SEQ ID NO:103); an RG2I polynucleotide sequence (SEQ ID NO:104) and its deduced polypeptide sequence (SEQ ID NO:105); an RG2J polynucleotide sequence (SEQ ID NO:106) and (SEQ ID NO:107), and its deduced polypeptide sequence (SEQ ID NO:108); an RG2K polynucleotide sequence (SEQ ID NO:109) and (SEQ ID NO:110), and its deduced polypeptide sequence (SEQ ID NO:111); an RG2L polynucleotide sequence (SEQ ID NO:112) and its deduced polypeptide sequence (SEQ ID NO:113); an RG2M polynucleotide sequence (SEQ ID NO:114) and its deduced polypeptide sequence (SEQ ID NO:115); an RG2N polynucleotide sequence (SEQ ID NO:116) and its deduced polypeptide sequence (SEQ ID NO:117); an RG2O

polynucleotide sequence (SEQ ID NO:118) and its deduced polypeptide sequence (SEQ ID NO:119); an RG2P polynucleotide sequence (SEQ ID NO:120) and its deduced polypeptide sequence (SEQ ID NO:121); an RG2Q polynucleotide sequence (SEQ ID NO:122) and its deduced polypeptide sequence (SEQ ID NO:123); RG2S polynucleotide sequence (SEQ ID NO:124) and its deduced polypeptide sequence (SEQ ID NO:125); an RG2T polynucleotide sequence (SEQ ID NO:126) and its deduced polypeptide sequence (SEQ ID NO:127); an RG2U polynucleotide sequence (SEQ ID NO:128) and its deduced polypeptide sequence (SEQ ID NO:129); and RG2V polynucleotide sequence (SEQ ID NO:130) and its deduced polypeptide sequence (SEQ ID NO:131); and, an RG2W polynucleotide sequence (SEQ ID NO:132) and its deduced polypeptide sequence (SEQ ID NO:133).

Characterization of New RG Family Groups and RG Species:

Further BAC insert characterization and sequencing, as discussed above, identified new RG polynucleotide sequences. The new sequences were characterized as belonging to new RG families; designated RG5 and RG7. These RG polynucleotides sequences, and their predicted translation products (the polypeptides which are encoded by these sequences) are summarized and listed below.

Identified and listed below is an RG5 family member, designated as the RG5 polynucleotide sequence set forth in SEQ ID NO:134, and its deduced polypeptide sequence (SEQ ID NO:135). This sequence contains an NBS region sequence.

Also identified and listed below is an RG7 family member, designated as the RG7 polynucleotide sequence set forth in SEQ ID NO:136. No deduced polypeptide sequence is given for the new RG7 family member as this sequence appears to be a pseudogene.

RG2A polynucleotide sequence (SEQ ID NO:87)

AAAGTTCATATCCAAGCTTGCCCTCCAAGCTAGCTCCTTCAATGGCACC
TCCTTCTCTTCAAAAGCACACAAGAAGCTTTCAAGCTCAACCACACTCA
CACAAGCTCTAGAACGAGGGTTAGGGCACATTTAGGGTTTTGCTCTCTGG
AAATGGTGTCTAAAAGTGAGGCCATAATGTTTCCTTATATAAGGCTCACTC
CCACAATTAGGCTTTCAATCTGAACGTANTACGCCAGTGACACTATGG
TACGCCCAACGTACTCGGTAGTCTCCGCGTCAANAATACACTCATGAGTA

CGCGCAACGTACTTTCCCTTACGCCAGCGTACTCAAAAGCCAAACATTC
TTTTCAAGGACTAATTTTGACAACCTTGAGGAAAGAAAAGGATCAAAGANA
TATACTTGAATTCGGGATGTTACAATGAAGTTGANACCTTGGCTAAAAA
ATTAAATTGGTTGTGGAAGCCGTTGGCTGAGCAAGCAACAAGGGTAAAAT
5 TCGTAATCTACAAATGGTGTATTCTATTTCTTCTTATTATTTTACTT
GATTTACGGGTAGTTTTTTTTTCTTACAAAAAATATTAAAGTTGATAAAG
TATAGCCACTAAAATTGACTTTTTTCCAAAACATAATGTCAAATGGTGCGT
ATATGTATCATGTTGTATTANATAATGAATATGATGATNCTGTTCTATTT
AANCCGAAAAAATTATCTAATGATTTTATATTGGAAAACAAAGTTGTGAT
10 TTTTNGCATAATATAATCAAATCCNCTTTTGTNTGGGAGGTGGATAAATG
TGGTAAATTTANAACAAGTGTTTTACNTTGAAGGGTNTGGAAAGGTGA
AAAAAGTTAAAATGATAAAATGTTTACACAAATGTTGTATCCGACTGAAT
ATNATGTTTAAGGATNATTGTATTAAATTGTTGATATATAGTAAGCATAA
ATATTTAGAATTGTGACTTAAATTTATAAGTTATNCNACTGGATTGAAA
15 CATTTTTGATATANATTAGGAATGAAAATGAGCAACCCTAACATACTTAT
CTTTGGTAGTTTGGTTATTATATTTTTATTANAATATAGAANCATCCCTT
TATTTTAAACCCATATTGTGGACGACTTGAATAAATGGGAAAAATGTAC
CTTGCTATTTAGCACAAAAAAATTATAAAAATGTACATTGCTATTTAGCA
CAAACAAAAAATAAACTTATCCTTTTTGCATTAGGTCACAAAGAAATA
20 TAAATGGGAAATGTGTTGCTATTTAATGCACTAAAAGAACTATTTTGC
CTTTATTAAACCGGGTAAACCAATAGAAAAATGGAAGTACATTGTCATTT
AGCATGAAAAAAATAACTTTCCATTTTTTGCATCCGGTCACAATAATAG
AAAAATGAAAGTACGTTGCTATTTAGCGAACTAACTTCCTTTTTTCTTT
TTGGCATCGTATCATAAAATATAGACTAAAATACGTTAGTTTTACATTTT
25 TAATACATTGAAATGTCTAATCCACATGTTATTCTATAAAAAGGGAAATG
TAATTTACTTATTCTTTGATTCTTTGGCTTCTTTTTAGTACCCAAAACAT
CCCTCTATCCATCTATTCCAATAAATAATGAAACTATATTCCTTCCA
TTGTAGGGATGTTATAAATTTTGTAATTGTTTTTATGCAAAAAAGTGTTT
TTTGTTAACTAGATTAACGAGATTCATTTTTTCAGCATTTTAGGAGAAGTT
30 CATCCATCTTTTGGATATGAAGTGCAAGCCAAGTTCTTTAACATGGAATA
TGAGGTCCCTATATGCTCAAAAAATAGCAAATGAGAAATTTTTTAAATTG
GATCCCCATAAAAGAAAATTTGTTAATGGTTGTTTAAATATTGGTCAATG
TGTCCACCGGATGAGCATAATACTAGTTTATAAGGGGTAAAGGTGGGTTT
GGTGGGCCCATTATCTTTATTATTTCTAAAAGTCAGAATTAAGTAAAAA
35 AAATTATAAGATAAATACCATAAGGATAAAAAATCATTTTATTTGGACCA
AAGACCAAAGTTGTTAAGGGGCTGTTTGTTTTTTTTTGTGAAGAGCTGTGC
AACCACTTTTGTCTGCGCCGCACAGACAACGTGCAGACATATGCCCTCGC
AGAGTGTTTGTTTTTTGAAGGTGCGCAGACCAAAAAACGTCTGCGCGAG
GTCATCCTGGCGCATATATGTGTCACTGTCTTCAAAGGTCTTCAGACCTC
40 ATTTTAAACCAAAAAAAGACCACCGGTTTTTTTTTTTTTTTTNTTC
TTTCTCTTGTAGCTGAAAATGCATTTTAAATCTTTATGACATGAAATTAA
GTTTGAAAAATTAATTTATTTCAACAGCTGTAGACGTAAAAACAAACAG
TCTTCTTGTTGCAGACTGTGGACATTTGGTCCACCTCTTACCGCAGAG

ACTTGCAGATGTGGTCCGCAGACTGCAGACATTTTGGCTTCAAATAAACA
AACATCACCTAATTTGACTACACCACACGGACCTCCAATGTAACAAAAAA
AAGGTTGAAACAAAGTTGCCTATTTCTCCATATCCAGGGGCCATTTATGT
AAGAGTTATCTAAATTTTAGTTTCGGTAGATCAGTTCTCACATTTTAACCG
5 GGTAAAGTGTATGTGTGTACGCGCGCACCTGAAAGGTTTGAANGTAACTT
CCAACTGAANCAANAATCGATATGAAGTATCAAGTTAGAGGTTCAATTG
GTGAAGGAATCAGCTGGAGGTTGGGGAATCGAGCTTCCACTATTAAGGTA
AAATCCATAACCCTAAATGTTGGTACGCTCATATATCAAATTGCGTGTTT
TGTTGAATGAAAAAAGCATGCTCAAAAAACCAGTGTAAGGCACGGTATAT
10 GACATATTTATAGTTACTGATAACAAATTATGATAATTTTGGGTTTACGT
AAGTTAGGATTCGTA CTTCAACCAAATGTAATAGTTTTTGTGAGTCTATC
TATGTATTTGGGGAATCACATTAGCAACGGGATTGTACTAGTAATTCGAA
AAAGTCTTTTAAATAATTTTTCTGTTTATAATTTATGAATAGTTTTAGCG
ACATCTAATATTAAATAGAATGTATCTGATATTGAATTAATGTCCTTAAT
15 GTG.AACATAGACCTTTTCCATTTACTAATGCCTAATTATTAGTTTCTAAT
CAATAAATTTTAATTTCTGTTTTATGCTTCTAAGACAATAAAAAATCCATG
ATTTACCTTTAAATATTAACAAAAATGACCATAAATAAATAAAAAATTAG
GAT.ACCAAACCCCCCGCCATGCCCAATGTCTAAATATTCTTGATGCTTT
TGCTTTTCCCTCTTTTCCTTGTTAGTCTATTATTCTGGAGAGTTTGAGAG
20 AGTTTCATACAAGAAAATTTCAAGAAGAAAGCAAAGGTCCAGGTATTCTC
TTTTCTTAATTATGTATTAACCTTACAAGCATTTTTTACACGATCCATGGT
TTTTTGTGTATGTTTTTCAAATTGAACTAGATTGGGACTTTTGCCCTTG
ATG.ATTCATAAGATATTGCATGGAGTTGAGATTGTGTAAGAAAAGTGGTG
AAT.AGAAAGAGCAAGTGAATCCAGATATAGTATTGGTAATATATGATGAT
25 GAG.ATAGAGATATGTTAAACTGGCTAGAAAATTGTTTTAATTTGAAATT
TAGGTTGTTGAATTTGAAAGATACCAAGCTAATAACTAATTAGTTATGCT
AAATAGTTATAAAGAACAACAACTCGTAGTTTTTTTTTTCATGATTTTCA
ACCTCTTCGTACCAAATAAATTATAACAAAATTGAATATCATTCTCTGC
AATCAATTTTAACTTTTGTATTATCATCATGTCTAAAATTGCCACAAGT
30 TTATTTTCATAGTCATATTGGATTATGAAAGGACTATTTTTACCAATTAC
ATCTTTACTTTATGGCCAAAGCTAATACAATCCGACTAACTAAAGGATT
CTAGGATGCATATAGTTTGCTCCCCGATTATAGATTTCTATCTAATTTGT
CTATTGTACTAATTTAGGTGCCACCACAAGTAAATTCCTGAAATGGATGT
CGTTAATGCCATTCTTAAACCAGTTGTCTGAGACTCTCATGGTACCCGTTA
35 AGAAACACATAGGGTACCTCATTTCTGTCAGGCAATATATGAGGGGAAATG
GGTATCAAAATGAGGGGATTGAATGCTACAAGACTTGGTGTCTGAAGAGCA
CGTGAACCGGAACATAAGCAACCAGCTTGAGGTTCCAGCCCAAGTCAGGG
GTTGGTTTGAAGAAGTAGGAAAGATCAATGCAAAAGTGGAATAATTCCT
AGCGATGTTGGCAGTTGTTTCAATCTTAAGGTTAGACACGGGGTCGAAA
40 GAGAGCCTCCAAGATAATTGAGGACATCGACAGTGTCATGAGAGAACACT
CTATCATCATTTGGAATGATCATTCCATTTCCTTTAGGAAGAATTGATTCC
ACGAAAGCATCCACCTCAATACCATCAACCGATCATCATGATGAGTTCCA
GTC.AAGAGAGCAAACCTTTCACAGAAGCACTAAACGCACTCGATCCTAACC

ACAAATCCCACATGATAGCCTTATGGGGAATGGGCGGAGTGGGGAAGACG
ACAATGATGCATCGGCTCAAAAAGGTTGTGAAAGAAAAGAAAATGTTTAA
TTTTATAATTGAGGCGGTTGTAGGGGAAAAAACAGACCCCATTGCTATTC
AATCAGCTGTAGCAGATTACCTAGGTATAGAGCTCAATGAAAAAACTAAA
5 CCAGCAAGAACTGAGAAGCTTCGGAAATGGTTTGTGGACAATTCTGGTGG
TAAGAAGATCCTAGTCATACTCGACGATGTATGGCAGTTTGTGGATCTGA
ATGATATTGGTTTAAGTCCTTTACCAAATCAAGGTGTCGACTTCAAGGTG
TTGTTGACATCACGAGACAAAGATGTTTGCCTGAGATGGGAGCTGAAGT
TAATTCAACTTTTAATGTGAAAATGTTAATAGAAACAGAAGCACAAAGTT
10 TATCCACCAATTTATAGAAATTCGGATGATGTTGATCCTGAGCTCCAT
AATATAGGAGTGAATATTGTAAGGAAGTGTGGGGGTCTACCCATTGCCAT
AAAAACCATGGCGTGTACTCTTAGAGGAAAAAGCAAGGATGCATGGAAGA
ATGCACTTCTTCGTTTAGAGCACTATGACATTGAAAATATTGTTAATGGA
GTTTTTAAAATGAGTTACGACAATCTCCAAGATGAGGAGACTAAATCCAC
15 CTTTTTGCTTTGTGGAATGTATCCCGAAGACTTTGATATTCTTACCGAGG
AGTTGGTGAGGTATGGATGGGGGTGAAATTATTTAAAAAAGTGTATACT
ATAGGAGAAGCAAGAACCAGGCTCAACACATGCATTGAGCGGCTCATTCA
TACAAATTTGTTGATGGAAGTTGATGATGTTAGGTGCATCAAGATGCATG
ATCTTGTTTCGTGCTTTTGTGTTTGGATATGTATTCTAAAGTCGAGCATGCT
20 TCCATTGTCAACCATAGTAATACTAGAGTGGCATGCAGATAATATGCA
CGACTCTTGTAAGAACTTTTCATTAACATGCAAGGGTATGTCTAAGTTTC
CTACAGACCTGAAGTTTCCAAACCTCTCCATTTTGAACTTATGCATGAA
GATATATCATTGAGGTTTCCCAAAAACCTTTTATGAAGAAATGGAGAAGCT
TGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCCTCATCAC
25 CTCAATGTTCCGTCAACCTTCGCGTGTTTCATCTACATAAATGCTCGTTA
GTGATGTTTGACTGCTCTTGTATTGGAAATCTGTCGAATCTAGAAGTGCT
TAGCTTTGCTGATTCTGCCATTGACCGGTTGCCTTCCACAATCGGAAAGT
TGAAGAAGCTAAGGCTACTGGATTTGACGAATTGTTATGGTGTTCGTATA
GATAATGGTGTCTTAAAAAAATTGGTCAAACCTGGAGGAGCTCTATATGAC
30 AGTGGTTGATCGAGGTCGAAAGGCGATTAGCCTCACAGATGATAACTGCA
AGGAGATGGCAGAGCGTTCAAAGATATTTATGCATTAGAACTTGAGTTC
TTTGAAAACGATGCTCAACCAAAGAATATGTCATTTGAGAAGCTACAACG
ATTCCAGATCTCAGTGGGGCGCTATTTATATGGAGATTCCATAAAGAGTA
GGCACTCGTATGAAAACACATTGAAGTTGGTTCTTGAAAAAGGTGAATTA
35 TTGGAAGCTCGAATGAACGAGTTGTTTAAGAAAACAGAGGTGTTATGTTT
AAGTGTGGGAGATATGAATGATCTTGAAGATATTGAGGTAAAGTCATCCT
CACAACTTCTTCAATCTTCTTCGTTCAACAATTTAAGAGTCCTTGTCGTT
TCAAAGTGTGCAGAGTTGAAACACTTCTTCACACCTGGTGTGCAACAC
TTTAAAAAAGCTTGAGCATCTTGAAGTTTACAAATGTGATAATATGGAAG
40 AACTCATACGTAGCAGGGGTAGTGAAGAAGAGACGATTACATTCCCAAG
CTGAAGTTTTTATCTTTGTGTGGGCTACCAAAGCTATCGGGTTTGTGCGA
TAATGTCAAAATAATTGAGCTACCACAACCTCATGGAGTTGGAAGTTGACG
ACATTCCAGGTTTCACAAGCATATATCCCATGAAAAAGTTTGAAACATTT

AGTTTGTGGAAGGAAGAGGTAAATATAAATTTTTAATGCTAATACATTAC
AAAGGATCTTTTCAGTTAAATCTTTCAAAATATATTGTAATTTGATTGTA
TGGGGTATTATTGTTGGATGGGACTATTAATAAATGATTATCTTGCAGGT
TCTGATTCTAAGTTAGAGAACTGCATGTTAGTAGTATGTGGAATCTGA
5 AGGAGATATGGCCTTGCGAATTTAATATGAGTGAGGAAGTTAAGTTCAGA
GAGATTAAAGTGAGTAACTGTGATAAGCTTGTGAATTTGTTTCCGCACAA
GCCCATATCTCTGCTGCATCATCTTGAAGAGCTTAAAGTCAAGAATTGTG
GTTCCATTGAATCGTTATTCAACATCCATTTGGATTGTGTTGGTGCAACT
GGAGATGAATACAACAACAGTGGTGTGAAGAATTATTAAAGTGATCAGTTG
10 TGATAAGCTTGTGAATCTCTTTCCACACAATCCCATGTCTATACTGCATC
ATCTTGAAGAGCTTGAAGTCGAGAATTGTGGTTCCATTGAATCGTTATTC
AACATTGACTTGGATTGTGCTGGTGCAATTGGGCAAGAAGACAACAGCAT
CAGCTTAAGAAACATCAAAGTGGAGAATTTAGGGAAGCTAAGAGAGGTGT
GGAGGATAAAAGGTGGAGATAACTCTCGTCCCCCTTGTTTCATGGCTTTCAA
15 TCTGTTGAAAGCATAAGGGTTACAAAATGTAAAGAAGTTTAGAAATGTATT
CACACCTACCACCACAAATTTTAATCTGGGGGCACTTTTGGAGATTTCAA
TAGATGACTGCGGAGAAAACAGGGGAAATGACGAATCGGAAGAGAGTAGC
CATGAGCAAGAGCAGGTAAGGATTTCAATTTCACTGTCTTAATTAATGAT
TAAGCTCCTGCTTTTTTGAATAAAAAAGGGACAAACCATTTTCATGACTTAA
20 TGTAGCAATACAAGTCATGTATAAGAGTGACCAACTCTTTTTTATTTATA
AAATGACTACAAAATATTTTTTTTCATTAGAGATCATGTATAAATGTGAC
TAATTTTTTCATCACCTAACTTTAGTTGATAAATCTTTATAAATGTCACTA
GTTACTTTTCAGTAAAATAACAAATTTAATAAATTATCAACAAAAAGCAT
CAACTAAAAAAATCCCACAACCCGTAATAATTTAAAAATAAAAGGATTTAA
25 CATCTAATACGAACAATTTTTTTTCTAAACATGATTTGGACCAAATATCA
CCAGCAACTCAAGTTTGGAAATCGATTCAGCTTAAACTTGACCAGCATAA
TTAGATAGATGAGAGTTGAAGCTAAAGTGCCTATATAAGTTCGTTTCATC
TTTTTTCTTGATCTTGATAGCAAGTTGAATGATTTTCTTCTTCAAAATTG
ATAAAAATCTACATTATAAAGAGACTAGCTTGAAAAAAAATGGTCTAGGT
30 GGGTCTTGGGTCTGTTAGATGAAGATGGAAGGGGAGAGTAGATTTCAAA
GACACAACACATCCTTCATTTTATTTATTTATTATTATTATTTTGTG
ATATCTTGCTCATATTTGTTACAGATATGTGAGGTCTATTAATCTTTTTA
AATATATAAAAAAATAAATAACATAAATGAGAAAATTAAATAAAGAATAA
ATT.AATAAGGGCACAATAGTCTTTTtaggTAAGACAAGGACCAAACACGC
35 AACAAAAATAAACAGTAGGGACCATCCGATTTAAAAAAAATAATTAGGGA
CCAAAAACATAAATTCCCCCAAACCATAGGGACCATTCATGTAATTTACT
CTTACTTTTCGTTTTGTTCATATTTGGGTAACTATTTTTTTTGTACACAT
CTAGGTAACGAACTTGTGAAGTGTTCCCATTTAGGATGTGACCTACTAC
AACCGATCATAATAGTCATATGTGAACACTTCCAACAACCTTTATTACTTA
40 GGTGTGTACAAAAAACAATAGTTACCATGATGTGAACATACTGAAAAAT
TAATTACCTTAGCAAGTTATTTTCCCATTTAGGTTGTATGGAAACAGTTC
CGTGAGACCGTGACTTGGATGGTAGATAAATTTAGTAACTTAACCCTTC
AATTAACCTACCTTTTTCTTATTAACCTCAATTTCAACCTAAATTCTGATT

CTTGTTTGAAAGTAAGTTGCATCTTTATTTTTGTATTATCTTGTTGCATA
GGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAAAGATCCAACTA
TTTTTAATCTGTTGGCATTTCATCATTTGCAACTGTTTCTTGAAAAA
AAATACCTAAAATCAAAATAACCATTTTCAAATCCAAAATTATAAGAGAG
5 AATTGTAAATGGACATGGAATCATAAATCATTAACACAGTTCAGTAAACA
AGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAAA
GAGACATTACAAGAAGCCACTGACAGTATTTCTAATGTTGTATTCCCATC
CTGTCTCATGCACTCTTTTCATAACCTCCAGAACTTATATTGAACAGAG
TTAAAGGAGTGGAGGTGGTGTGTTGAGATAGAGAGTGAGAGTCCAACAAGT
10 AGAGAATTGGTAACAACCTACCATAACCAACAACAACCTATTATACTTCC
CAACCTCCAGGAATTGATTCTATGGAATATGGACAACATGAGTCATGTGT
GGAAGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCA
GAATCCCCATTCCACAACCTCACAACCATAAAAATTATGTATTGCAAAAG
CATTAAGTACTTGTTTTCGCCTCTCATGGCAGAACTTCTTTCCAACCTAA
15 AGCATATCAAGATAAGAGAGTGTGATGGTATTGGAGAAGTTGTTTCAAAC
AGAGATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACACAAC
CACCACCTTGTTCCCTAGTCTTGATTCTCTCACTCTAAGTTTCCTGGAGA
ATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAA
ATATCTTTCAATAATACCACTGCAACTACTGCTGTTCTTGATCAATTTGA
20 GGTATGCTTTGTACATATTCAATTATTTATTTAATTTCCTTTTTTATTTG
CAATATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAAT
TACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAACTCT
AGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGAT
GGGTAATATAGGCAATTTAAGTTTATTTCTGTAAAGCAGTATTTAGCA
25 AGTACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGTC
ATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATCAGGGGTCAT
CAGGTGACAGATATTGTAGAATAGAACAATATATAATATCACCCAAAAC
ATTTTTTCTAAGGTTATTCTGTAAATATGTGCTTTCTTGTTTTTCATNGA
ATTNGCATTCGTATATTTTAGGTGTAAAGTGATTTTNTCTTCAATAAAT
30 CCCGAAATTAATTAACAAAAACAAAAGTACATTTTGTATGTGGAG
AGCACTGGTATCACTTAGTATATAAAAAGCTTGATTTTGAATTAACCTTC
TTATACAAAAGTTGTGTATATAGTTTAATTAGTTTACATCATTTTCCA
TGTGGTGTGTCAGTTGTCTGAAGCAGGTGGTGTGTTCTTGAGCTTATGCC
AATACGCTAGAGAGATGAGAATAGAATTCTGCAATGCATTGTCAAGTGTA
35 ATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGACAGT
AAGTGATTGCAAAGGGATGAAGGAGGTATTTGAACTCAATTAAGGAGGA
GCAGCAACAAAAACAACAAGAGTGGTGCAGGTGAGGAAGGAATTCCAAGA
GTAAATAACAATGTTATTATGCTTTCTGGTCTGAAGATATTGGAAATCAG
CTTTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGAAAGCC
40 TGAGACAGCTCCAAGAGTTAAAGATAACATTTTGCTACGGAATGAAAGTG
ATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGTAAACAACAACAAC
AACACAATAACGAAGGGGGCATCATCATCATCTTCTTCATCTTCTA
AGGAGGTTGTGGTCTTTCCTCGTCTCAAATCCATTGAACTAAATGATGTA

CCAGAGCTGGTAGGATTCTTCTTGGGGAAGAATGAGTTCCGGTTGCCTTC
ATTGGAAGAAGTTACCATCAAGTATTGCTCAAAAATGATGGTGTTCAG
CTGGTGGGTCCACAGCTCCCCAACTCAAGTATATACACACAGAATTAGGC
AGACATGCTCTTGATCAAGAATCTGGCCTTAACCTTCATCAGGTATATAT
5 ATTTCTTTAATTGGCATCATCTAATTAAGAAAGATATCATTCCTGCCAAG
TAAATTTACTTCAAACACATTCACACTGGTTTCAGTCTAAGTTTATGTTG
TTCTAGGAAGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTT
CAGTGGAAAGGGTATTTTAGGTATTTTCTGTCAAAAGTTGTTATTGCAGG
CTTTTATAGTACCTGGAATCGTGTGTGGGAGGAGCATTATTATTCTGATTT
10 GCTTGTTTCTTTATCATTTTTTCTTAGCCTCTGGAACAGCTAGAAACCCT
TTT.AATCTTTTGATTTTCAATGACAAAATTTTTCCTGTTACTACATTTGA
TTGTTGTTCTTCATGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCT
TTTGATTGTTATTTTCATATCATGTTAGTCACTTGAATCAAGCTTTTCTA
TTTTCAACCAGGGCAAAAGGTCAAAAGTAACCTACTTTATGAGATCAAAA
15 ACAGCAACCCATCGGATAACTTTTAGTTGGAGTTAATAGTTACAATTACC
ATTGTGATTAATAATTATAATATCTTGTATTAATTCATAAAAATTGGTAC
AGCACATATATGACATTTCAAAGGTTTTTGTGTTGACATATATATGCCTCT
GGCGTTTTCTTTATTGGACATGCAGACCTCATTCCAAAGTTTATACGGTG
ACACCTTGGGCCCTGTAACCTCAGAAGGGACAACCTTGTTCTTTTCATAAC
20 TTG.ATCGAATTATATATGGAATTTAATGATGCTGTTAAAAAGATTATTCC
ATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAGATTCATGTGACTT
ATTGTAATTGGGTAGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGG
AGAAATGGAAATAGTGGAATTGGTTTTGATGAATCGTCACAAACAACACTAC
CACTACTCTTGTCAATCTTCCAAACCTCAGAGAAATGAAGTTATGGTATC
25 TAAATTGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTTGAG
TTTCCAAACCTAACAAGAGTCGATATATGGGGATGTGATAGGTTAGAACA
TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAC
GCATATGGAAGTGCAGTCAGATAGAGGTCGTGATTGTTTCAGGATGCAGAT
GTTTGTGTAGAAGAAGACAAAGAGAAAGAATCTGATGGCAAGACGAATAA
30 GGAGATACTTGTGTTACCTCGTCTAAAGTCCTTGATATTAAACACCTTC
CAWGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTA
TTGGATACYTTGGAAATCTACRAATGCCCAGCAATAACCACCTTCACCAA
GGGAAATTCRCTACTCCACAGCTAAAAGAAATTGAAACAMATTTTGGCT
TCTTTTATGCTGCAGGGGAAAAAGACATCAACTCCTCTATTATAAAGATC
35 AAACAACAGGTAAACCAGATCTTTGTTGCTTNNATAATTCTTAAACNACA
TNTGAAAAGCTTCATGCAAGTTTTTTTTNGTTATATNGTCAAAAACCGCAA
CCTACATTTTCAGCTTTANATTTATGTACTTTATGCAGGATTTCAAACAA
GACTCTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTTCATGT
TCCTAGTNGCCTATTAATTAAGGCCTTTTAGTTTCGNGATTTTGGATGT
40 ATTCTTCATGATGATGTCAATCTTCTAATAACCCATTTCATTGTTTGGTTG
AATGTTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTTCATCA
TATGAAGGACATTAAAGAACATGGATGCTCTGAAGATGTTGGGAACACA

RG2A deduced polypeptide sequence (SEQ ID NO:88)

MDVVNAILKPVVETLMVPVKKHIGYLISCRQYMREMGIMRGLNATRLGVEEHVN
RNISNQLEVPAQVRGWFEVVGKINAKVENFPSDVGSCFNLKVRHGVGKRASKIIEDI
DSVMREHSIIWNDHSIPLGRIDSTKASTSIPSTDHHDEFQSREQTFTEALNALDPNHK
5 SHMIALWGMGGVGKTTMMHRLKKVVKEKKMFNFIEAVVGEKTDPIAIQSAVADY
LGIELNEKTKPARTEKLRKWFVDNSGGKKILVILDDVWQFVDLNDIGLSPLPNQGV
DFKVLLTSRDKDVCTEMGAEVNSTFNVKMLIETEAQSLFHQFIEISDDVDPELHNIG
VNIVRKCGGLPIAKTMACTLRGKSKDAWKNALLRLEHYDIENIVNGVFKMSYDNL
QDEETKSTFLLCGMYPEDFDILTEELVRYGWGLKLFKKVYTIGEARTLNTCIERLI
10 HTNLLMEVDDVRCIKMHDLVRAFVLDMYSKVEHASIVNHSNTLEWHADNMHDSC
KRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFKPNFYEEMEKLEVISYDKMKY
PLLSSPQCSVNLRVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSAIDRLPSTIGKLL
KLRLDLTNCYGVRIIDNGVLKKLVKLEELYMTVVDRGRKAISLTDDNCKEMAERS
KDIYALELEFFENDAQPKNMSFEKLQRFQISVGRYLYGDSIKSRHSYENTLKLVLK
15 GELLEARMNELFKKTEVLCLSVGDMNDLEDIEVKSSSQLLQSSSFNNLRVLVSKC
AELKHFFTPGVANTLKKLEHLEVYKCDNMEELIRSRGSEEETITFPKLLKFLSLCGLP
KLSGLCDNVKIIELPQLMELELDDIPGFTSIYPMKKFETFSLLKEEVLPKLEKLHVSS
MWNLKEIWPCEFNMSSEVKFREIKVSNCDKLVNLFPHKPISSLHHLEELKVKNCGSI
ESLFINHLDCVGATGDEYNNSGVRIKVISCDKLVNLFPHNPMSILHHLEEELEVEN
20 GSIESLNFNIDLCAGAIQEDNSISLRNIKVENLGKLREVVRIKGGDNSRPLVHGFQS
VESIRVTCKKKFRNVFTPTTTNFNLGALLEISIDDCGENRGNDSEESSHEQEIEILS
EKETLQEATDSISNVVFPSCLMHSFHNLQKLILNRVKGVEVVFEIESESPTSRELVT
HHNQQQPIILPNLQELILWNMDNMSHVWKCNSWNKFFTLPKQQSES PFHNLTTIKI
MYCKSIKYLFSPLMAELLSNLKHIKIRECDGIGEVVSNRDEDEEMTTFTSTHTTTT
25 LFPSLDSLTLFLENLKCIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVSW
SLCQYAREMRIEFCNALSSVIPCYAAGQMOKLQVLTVSDCKGMKEVFETQLRRSSN
KNNKSGAGEEGIPRVNNNVIMLSGLKILEISFCGGLEHIFTFSALESRLQLQELKITFC
YGMKVIVKKEEDEYGEQ.TTTTTTITKGASSSSSSSSSKEVVVFPRLKSIELNDVPELV
GFFLGKNEFRLPSLEEVTIKYCSKMMVFAAGGSTAPQLKYIHTELGRHALDQESGL
30 NFHQTSFQSLYGDTLGPVTSEGTTCFSFHNLIELYMEFNDAVKKIIPSELLQLQKLEK
IHVTYCNWVEEVFETALEAAGRNGNSGIGFDESSQTTTTTLVNLPNLREMKLWYL
NCLRYIWKSQWTAFEFPNLTRVDIWGCDRLEHVFTSSMVGSLLQLQELRIWNCSQ
IEVTVQDADVCVEEDKEKESDGKTNKEILVLPRLKSLILKHLPCCLKGFSLGKEDFSF
PLDITLEIYKCPAITTTFTKGNSTTPQLKEIETHFGFFYAAGEKDINSSI KIKQQDFKQ
35 DSD.CEVNIK

RG2B polynucleotide sequence (SEQ ID NO:89)

TTTTTTAAGATCAGGGATTCAAATTCAGCCCTAGTGATTACAATTGTGTC
TAAACTTTCCCATACCTTCACATTATTGTAAGTATACTTTCTCAGTTTCT
40 CTCTTGGAAGCTTCCTTGGTATTTTAACTCGTGTTCTAATATTAACTCT
GATAGTTATTTTGGCCAATCTACTATCTGCATGTCCGGTTATTGAATCCG
AAGGCACTGGAATCTTGGATTCCATTCCGTTGTGTGTTTGGTTGCCAAAT

GAACGGAATTGAATTATGTAAGATTCCTTCAAAAATCCATGTTTAGGTATA
TCGTTGTTTCTTGGGATGGATGGTAAAGAACGGAATTTCTCCTGTTTCATT
TTTTAATGAAAGACCAAATTGACCTTATAAACCTGTAAAAAAATTACAT
TCCAGTTTTCTTAACAACTGAAAATGGTAAAGGAGTGTGATTGAATTCC
5 AATCTGTTTCCTGTCCAAAACACGTGACGGAATATTACAATTCCTTCAA
TTTCATTTTCTTAAATTGTTATTCCCTTTCTTACAAAAACAAGGTAAACG
AAACACCCGCTTACTTAATCATACTCCTACATGATGTAAATGAAAAGGGT
ATAAATGGTATTTTATTCACAGGGATGAGTCACCATGGTCATGAAAGAAT
CATTAACCGCCCTTACCCAATTCATGTTTGCCCCATAAATATGATTTAAA
10 GTAATATTGGCTTATGGGATTCAAGTTGACTTTTTTGTGGCGAAGAAATA
ATGAAAATCTTCATTTCTAAAGTGTCTTCTACCACTGACATTTTCTAAGA
AAGAACTTGCTAGAAGAAGGTGGGTGTTTAGTCTTTTTACTCTTTAAAT
GTGAAGACTGTTGAGTTATTATTATTTTGGCAACTATGGACAACCTG
TTTAGTTTTTTTTTTTCCCAATATCCATTTATATGCGATTTATTTCTGA
15 AATAATTTTATCAAAACGCAGGAAACAATGTAGAATAATACTGGTATAAT
TAATTATATAAAGTTATTAGGCTGAAATCTTGAGGCTACTATAATTTAAT
TATCATAATTTGAAAATCATCAAATTGTATTCCATGTATATTTATGTTAT
CAGATAATTAATAATATGTGAGCCACACAAATCCACATCATCAGACACCC
CACCTTATTGTCGGCTACCTCACCACTTGCATGATCCCGACATCTTCCCA
20 ACCCCACCGACGACTTGGGGTCTCCTTAATATATCAATTATTTTCTGTAA
GTATTTATTTGTGTAAATGTGTAATGTCATTTTACCTTTTTTCTAATATA
TACAGAAACATAAATTTTAAATGAAATTCAACTGCGTTTCATTCTTGCAT
TAA.AAAAAAAGACTGTACTGTTGTCAATATTTTACTTATAACCTGATTAA
TTAATTAAAGCGTAATTGCATAATTTGCATTAGGTTGTAATTTTGTGTTT
25 TATAGGGAGGGTGAGGGTCACCGGGAATCAAAGCACTTATGTAAAAGCAG
GGA.AATACAAAAAATTTACTCGAAACAAATTTTATTCAATTTAAGTGAGA
TAATAATGTTCTGATTAGATTATGAGAACTAGGAGATTTAAGTGATATAT
CCC.ATTTAAAAGAAATTGCATTATTAATTTTGGATCTCTTGATGATGACA
AAATTAACCTCGTGACAGGTTATATATCATATACAAAATGAGTGGCTATGC
30 TTTCGCTTTCCAAAAAGCAATTATAGTTATACTACACCTACAAATTTTAA
AAGGGGTAAACATATCAAAATACTTGATAAGTAATTATATAAATATGCA
TTT.AACCCTCTAAAGAAAATGCTACTAAGCTTGGACCATCTCAGAATTAC
AATCATACCCTTCCCCCTCAAAAAAGATTTCGTATATATCATGTCATTTGGC
ATTCATTTCTTTTTCACAATTCATAGTTCTATTCTCAAAAAATTCGAGTT
35 CTCGTATTTGTAAGGAAGATCAGAAGAGACTGTTACACAGGTACTCTCT
TTT.ATTTATTGATTCACATTCATATATGTTATTGTTTTCTTGCTTAATGG
TTTCGTCACTCTAACTGCGCTTGCTGATTTAAATTTCTTCACTTTCTTCC
ACGGATTTTTTAAATATTAGTTTTGTGAATGAACAATTGGTGAAGGAAAG
AAACATGGGAGTCTTTTCTAAAGTAAACCTAGATACTTAGGTTATAAGGG
40 TATATGCTAAAATGAACTATGCCCATTCACCTTTGCCTTTTCTTTTACTT
TTT.AGTTTTTAGAATCCAAGTTTTTCATATGTATCTCGATGTGTGAGAAGA
ATAGGCATTAGAAAGGTAAAGGACGTACATAAAATTGATTAATTAGTGAA
TGTTCTTTGATATCATTATTTTTACTCTCATAAAAAGCATATAGATCAAA

CACAAATTGCTACTTGTTAGTGTAACAACTTCGACTTAATAATGTTAATA
ATCAAGATTCTCTTGATTTCAACTATTTTCTAACCGAACAAGCTCACTAA
AAACTCATATTGCTTTGAGTCTGAGTGGTTTATATTTGGGGTTTTACATT
TAATTTTTTGTGCATGAATGTGAAAATAGACTGCTTATTGATTCTTTGTG
5 TTTTATTGAGTTGATTTTCATTATTACTACCTTACAAATTGCTCAGTGAT
AGATTTCCATTAATTTGCTAATTCGGTTGCTTCTAAATATGTAGGAGCTA
CTA.AAAGCAAAAATATCGAGCAATGTCGGACCCAACGGGGATTGCTGGTG
CCATTATTAACCCAATTGCTCAGACGGCCTTGGTTCCCGTTACGGACCAT
GTAGGCTACATGATTTTCTGCAGAAAATATGTGAGGGTCATGCAGATGAA
10 AATGACAGAGTTGAATACCTCAAGAATCAGTGTAGAGGAACACATTAGCC
GGAACACAAGAAATCATCTTCAGATTCCATCTCAAATAAGGAATGGTTG
GACCAAGTAGAAGGGATCAGAGCAAATGTGGAAAACCTTCCGATTGATGT
CATCACTTGTTGTAGTCTCAGGATCAGGCACAAGCTTGGACAGAAAGCCT
TCAAGATAACTGAGCAGATTGAAAGTCTAACGAGACAACTCTCCCTGATC
15 AGTTGGACTGATGATCCAGTTCCTCTAGGAAGAGTTGGTTCCATGAATGC
ATCCACCTCTGCATCATTAAGTGATGATTTCCCATCAAGAGAGAAAACCT
TTACACAAGCACTAAAAGCACTCGAACCCAACCAAAAATTCCACATGGTA
GCC TTGTGTGGGATGGGTGGAGTGGGGAAGACTAGAATGATGCAAAGGCT
GAAGAAGGCTGCTGAAGAAAAGAAATTGTTTAATTATATTGTTGGGGCAG
20 TTATAGGGGAAAAGACGGACCCCTTTGCCATTCAAGAAGCTATAGCAGAT
TACCTCGGTATACAACTCAATGAAAAAACTAAGCCAGCAAGAGCTGATAA
GCTTCGTGAATGGTTCAAAAAGAATTCAGATGGAGGTAAGACTAAGTTCC
TCATAGTACTTGACGATGTTTGGCAATTAGTTGATCTTGAAGATATTGGG
TTAAGTCCTTTTCCAAATCAAGGTGTCGACTTCAAGGTCTTGTTGACATC
25 ACGAGACTCACAAGTTTGCATATGATGGGGGTGAAGCTAATTCAATTA
TTAACGTGGGCCTTCTAACTGAAGCAGAAGCTCAAAGTCTGTTCCAACAA
TTTGTAGAACTTCTGAGCCCGAGCTCCAGAAGATAGGAGAGGATATCGT
AAGGAAGTGTTGCGGTCTACCTATTGCCATAAAAACCATGGCATGTACTC
TTAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCGCGCATAGAG
30 CACTATGACATTCACAATGTTGCGCCCAAAGTCTTTGAAACGAGCTACCA
CAATCTCCAAGAAGAGGAGACTAAATCCACTTTTTTAATGTGTGGTTTGT
TTCCCGAAGACTTCGATATTCCTACTGAGGAGTTGATGAGGTATGGATGG
GGCTTGAAGCTATTTGATAGAGTTTATACGATTAGAGAAGCAAGAACCAG
GCTCAACACCTGCATTGAGCGACTGGTGCAGACAAATTTGTTAATTGAAA
35 GTGATGATGTTGGGTGTGTCAAGATGCATGATCTGGTCCGTGCTTTTGT
TTGGGTATGTTTTCTGAAGTCGAGCATGCTTCTATTGTCAACCATGGTAA
TATGCCTGGGTGGCCTGATGAAAATGATATGATCGTGCCTCTTGCAAAA
GAATTTTATTAAACATGCAAGGGTATGATTGAGATTCCAGTAGACCTCAAG
TTTCCTAACTAACGATTTTGAACTTATGCATGGAGATAAGTCGCTAAG
40 GTTTCCTCAAGACTTTTATGAAGGAATGGAAAAGCTCCATGTTATATCAT
ACGATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCGATGCTCCACC
AACATTCGGGTGCTTCATCTCACTGAATGTTTATTAAAGATGTTTGATTG
CTCTTCTATCGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCAAATT

CTCACATTGAATGGTTACCTTCCACAGTCAGAAATTTAAAGAAGCTAAGG
TTACTTGATCTGAGATTTTGTGATGGTCTCCGTATAGAACAGGGTGTCTT
GAAAAGTTTTGTCAAACCTGAAGAATTTTATATTGGAGATGCATCTGGGT
TTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTTACAACCTTTCT
5 GCATTAGAATTCGCGTTCTTTAATAACAAGGCTGAAGTGAAAAATATGTC
ATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGATGCTCTTTTGATG
AAAATATCAATATGAGTAGCCACTCATACGAAAACATGTTGCAATTGGTG
ACCAACAAAGGTGATGTATTAGACTCTAACTTAATGGGTTATTTTTGAA
AACAGAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGATCTTGAAGATG
10 TTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCATTCTGCAATTA
AAAGTTCTTATTATTTCAAAGTGTGTAGAGTTGAGATACCTTTTCAAAC
CAATCTTGCAAACACTTTGTCAAGACTTGAGCATCTAGAAGTTTGTGAAT
GTGAGAATATGGAAGAACTCATACATACTGGAATTGGGGGTTGTGGAGAA
GAGACAATTACTTTCCCTAAGCTGAAGTTTTTATCTTTGAGTCAACTACC
15 GAAGTTATCAAGTTTGTGCCATAATGTCAACATAATTGGGCTACCACATC
TCGTAGACTTGATACTTAAGGGCATTCCAGGTTTCACAGTCATTTATCCG
CAGAACAAGTTGCGAACATCTAGTTTGTGTAAGGAAGGGGTAGATATATG
TTCTTTATGTTAATAACAATTTAAATAATATTTTCAACCAAATTTTCATAA
TATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGGCTATT
20 AAGGGATGATTATTTTGCAGGTGTGATTCCTAAGTTGGAGACACTTCAA
ATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACTTAGTGG
AGGTGAGAAAGTTAAGTTGAGAGCGATTAAAGTGAGTAGCTGTGATAAGC
TTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATCTTGAA
GAGCTTACAGTCGAGAATTGCGGTTCCATTGAGTCGTTATTCAACATTGA
25 CTTGGATTGTGTCGGTGCAATTGGAGAAGAAGACAACAAGAGCCTCTTAA
GAAGCATCAACGTGGAGAATTTAGGGAAGCTAAGAGAGGTGTGGAGGATA
AAAGGTGCAGATAACTCTCATCTCATCAACGGTTTTCAAGCTGTTGAAAG
CATAAAGATTGAAAAATGTAAGAGGTTTAGAAATATATTCACACCTATCA
CCGCCAATTTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAGGTTGC
30 GGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTCAATTTCACTTTC
TTAATTAATTAAGGACTAAGCTCCTGTTTTTTGAATAATAAAGAGGTGGG
ATGACTAACTTGGGCATCACAATTGCAACAAAATGTTACAAACCATGAA
ACGTTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTTAAAAA
TATGGCTTAAATTTTTTTTATATTTATGTATCAACATGATTTTTCATTAG
35 AGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATCAGAACTAATT
CTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTATGTG
AAATAAGCAACGGATTTAATAAGTTAACAACCTTAAATGTCATTTCTTAAC
AAAAAAAATATTTGGTTCAGAAAAACCGTAATTCAAGATAACTAAAATA
AAAATATTTGACATTCACCTAAGAGCATTTTTTTTTCTAAATATGATTGCA
40 AATGAATAAACTTAAATTTATACAGAAAATTCTTTTATATATGTTATAC
AAAATTTACAAATTGAAATTGGATATGTTAATTAACGGTTTATAATTCTG
GTATCACAAAGGGATATATAATAAAATATTATTTCTGTAGTCATTTGTA
ATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAGTTAA

ACTTTCATAATAAAAAATTTATAATTATTATTTATTTTAAATAAAATTATTA
ATTAAGAGATATATCAAAAATTTAAAGTTATTATAACTTCAAATTTAACA
TATAATTAGAAAATATATGATCATAACTTTCTGCAACTCTTCTTTTGTAT
TAAAATGACCAGAGAAGCTCTTAGTATATTTCTAATCAAAGTCTCAAAMC
5 TAATGAAGCATATAATTTGTGAAAATCAATTAGCATTAGGTTTTAAGAGT
CACCAAATTCAAAGAATAATCCAATGCTTTCATTACCACTATGGAGAAAA
TATTTTCTTAGTTTAAATGAAATGAAAACAAACATTCAAACATAATTGTTG
CTTATTAAACCAAAGACCCATTACTTAGCCAAGAGTTTAACAAAAAAAAAA
TTACATTCATGTATCATTATTCATGACTAGATATATATGAACATGAAGGG
10 AGTTTTTTATAGAAAATATAATCATAGATATTCAACATAACTTCAGGGAAT
TCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCAACCAAAG
AGAAGTTTAGCCTAGCATGGCTAAACTCAAGAACTAAAATAAGGATTAG
AAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTTGTTCTTG
ATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAATGCAAAGG
15 AGAGCCATTAAATTCGTATGTATTGATCCCTTCAAAGCTGCACCAACCT
CCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCTGAAGGACC
CTATGTGGGTGCCTTGCGCGGGTGGAGCTGCATACGAAAGGTCTTTGGTC
TTTGTGAGGGTGATGTTGTGCGGGATAGCTTGTCGCATGCTTCCGCGCGG
TTCACGCACATGTGCACAGGTGATGCATGGTGTGTGCGTTCTTGAGTTTT
20 GAGCCTCCGATGCTTAGTCCACTTGGCCCAATTCGAGTCCAATCAGCTTA
TAACCCATTTTTCTTCAAGTTATCTTCAAGTTAAGCCCAATTTGCCTTCT
CCAATCATCCATAACTTCACAGAATCGCCCGTTCATCTTAATCCCGGAT
GCACAATTATTCTCCCGTCTTCATTTTAAGCAAGATACCACCTTCTTCAT
GCTTCATCCATCAATAGTACACTTCATGTATCATCTCTACTAGTTATTTA
25 GTCCACAATCCTTGTTGTCTTCCAAATTTAATTATCTCATTTAGTTCCCG
TTCCGCTAGTTTCCTTAAAATTTGCAATTAAGCTCAGAGAAATATTAAGT
ACCCGAAATGGTCATAAAATAACAAAAAGGAAAATATGCATGAAGATTAA
CTAAATGATGAACGAAATATGCTAAAATAGACTATAAAATGAAGTAAATA
AAATGAAATTATCGCACTCCGACCACCCTTATAGGCTTGTAGTCCACCCA
30 CCCTTCATTCTTGTACCAATATGGGATGGAAACATCATTAAATTAAGCCA
AAAAGCTAACATATAAGGGTTTAGTGACAAAGGTAAGTACTAAAGATGAA
AATAATCCATTTTTCTTGTATATACACAACACACACATAGGGGCGAGACGT
AGGATTTCAAAGTACAGATTGTTGGTGGCACATAAGTGTTGCTGGTGACA
TTTTTTTTTTTTTTTTTACGTAGTGGCACAACAGTAGGAAAAACGAAAAAT
35 TCGAAATTTTTTACAATTTGTCTAAAAAAAACAGTGGTTGTTGGTGCCAC
TATGGACACCAAAGTTGAACTGCCCCACGCGCGCACACACACACACACA
CATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAAGAAAGAGAGAGA
GAGAGTTTGGGATGTGATACTTCTTTTAGGAAAATGGAGTTATATCTTTG
ATATTGTATTTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTATA
40 AGTTTTATTTATTTTGATATGAAAAAAAAGTCTTTTATACATTGGATTT
AACATAAAAATCCAACAATATTAATCAAAAAGACCAMACATGTGGACAMW
TATGTATATAAWTAATTCACAATAGTCTTTAGGAATAGNATTATATATAT
AATTAATTCTCAATGGTCTTAGGAATAGTAAGTTCTTATATTTCAAACCTT

TNGCCACAATTCTTTGKTTACTTWGACACTTYCCTCTCTCTAATTATATA
TATATATATATATATATATATATATATACACACACACACACACACTAG
ATGTGTGCCCGCGCAAAGCAGTGACGTNNNGGAGAANACTTTCTTAAGCA
TAAATAATTATTATATTTTTTTATTGGGTATTATATAATAAAAAATTACAA
5 CTTTTAAATAAAATATTTATGTTTATACTTTATATTTATATTGCTTGTAT
ACTATTAATATAATAAAATTAATATTTATGTCTAATTTATGAAATGTAAAT
TAATTTAAATACATGAATTTAATATTTTTTAAATTTTCAGTTTGCTTCAA
ATTGAGTTTCTTAATTATTTTTTTTAATTCANGTATTCAAACCTTTTGGTA
AGTATTAAAGAATTATTTATGCATAATTGATTTATACAAAAAATTTTGTA
10 ACTTATACATCTTAAAATTCAAGATATAACTAACATGTTTTACAATATAT
AT
TAAAGCGCAAAGGTCATAGGAATAGAATATTTTCTATTATTCTACGTTTT
GCCACAAAAGTTTGAACACTTTGCCACTTTTTGTCCCTCCTTAACCTTTT
CAATGTTTTGCGACAAAAGTTCCAAAATTTGCCACTTTGATCATTCCCTC
15 AACTTTTCACCGCAATTAGTTTGTGGAGTTGGCAGTTTTGATCCCCCTAA
CTTCGATATTCTCTACTGCTAGCCAAAAAGGGTTCCAGAGTTTCACACTT
TTGGTCCCTGACAGTAACCAAATGTGAGATGTCAAATTTTTGCCACATTA
GTTTGTGGAGTTGTCCCTTTTGGTCCCCCACATTTCGATATTCTANTATA
CGACCTTATTTTTNTCAAATAACAACACGTATATTTAATTACCAATTATA
20 GAAATAGATATCAAATAAAGTATTTGTAACACTGTGTAAGAACGGTGCTA
CTATAGGTAAAAATAAACATTTCAAAGTACGATATCCTAATTGGAAAAAG
AGTTTTAAAAAAATAACGACTAGGGGCGAGTTTTTTTTTACAAGTTTGTAT
CAAATCATATCAAATTTAAGGTGGAACGGTGACCACATTAACCAGAAAT
GTAATTTATTCTTTGATTTTGATAATTTTAAATATTTTGTGTGATCTAT
25 GTATTTAAAAGTAAACAACAAGAACAATAATCCAAAACCCTAAATTGCAA
GTCTCGCCCAATTTCTCTATCACTAGTCCTCACTTACGATGGCGTTACGT
CGCTCTCTCACTGCTTACAACCCTTTGTTGCTACTCATTACAATAACGAA
AAGTTGAATATCCATATATTTATTTGGATGTGGAATTGAACGAATCTCGT
CAAATTTTGTATTTTGTGATGGATTTGAGTAGAAGTTTGGGCAGAACGG
30 GAATGATGGTCTGCAAGTGGTTATAAACTTGATTCTGAGTTATTACTATA
TATGTAGCCTCTTTACAACGACCAAGGTTTCTTCCAGGTACCATTGATC
TTTTTAGAACTTAGTTTTCTGAAACACCCTGATTTGGATCAAATATCACC
AACAACTCTTAAAACTTGATTAATCAATTGTTTTCTTCATCTTGATAAC
AAGTGGAATGATTTTCTACTTAGATTAACCTGAAAAAAAAGGTCCATGTG
35 CGTCTGGTGGATCTGGTAAATGAAGATGGAAGGGAGAGCTGACTTTAAAG
ACACAAACACGTCACCATATCTCTTATTTTATTTTAAATTTGCTTTTGGT
GTATTTTCTTTTTTCTTATTTCTTTCTTTCTTGATCTCCAGATGGTATGT
GGTGTGGATAATTTACACCTAGAGATTGGGAACGATGGGAAGGGGTCTGT
GATTTATGGCTGGCCGAGTTTTACTTATTAACCTCAATTTCAACCTAAATT
40 CTGATTCTTGTTTGAAAATAAGTTGCATCTTTATTTTTGTATTATCTTGT
TGCATAGGATCCTTAGCATCTTTTAAATAATTTATTTGAAGGTGAAAGATC
CAACTATTTTTTAGCTGTTGGCATTTCATCATTTGCAACTGTTTCTTG
AAAAAAAATACCTAAAATAAAAAATAACCATTTTCAAATCCAAAATTATA

AGAGAGAATTGTAAATGGACATGGAATCATAAATCATTAACACAGTTCAG
TAAACAAGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCA
GAGAAAGAGACATTACAAGAAGCCACTGGCAGTATTTCAAATCTTGTATT
CCCATCCTGTCTCATGCACTCTTTTCATAACCTCCGTGTGCTTACATTGG
5 ATAATTATGAAGGAGTGGAGGTGGTATTTGAGATAGAGAGTGAGAGTCCA
ACATGTAGAGAATTGGTAACAACCTCGCAATAACCAACAACAGCCTATTAT
ACTTCCCTACCTCCAGGATTTGTATCTAAGGAATATGGACAACACGAGTC
ATGTGTGGAAGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAA
CAATCAGAATCCCCATTCCACAACCTCACAACCATAAATATTCTTAAATG
10 CAAAAGCATTAAGTACTTGTTCGCTCTCATGGCAGAACTTCTTTCCA
ACCTAAAGGATATCCGGATAAGTGAGTGTGATGGTATTAAAGAAGTTGTT
TCAAACAGAGATGATGAGGATGAAGAAATGACTACATTTACATCTACCCA
CACAACCACCACTTTGTTCCCTAGTCTTGATTCTCTCACTCTAAGTTTCC
TGGAGAATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGC
15 AATGAAATATCTTTCAATAATACCACTGCAACTACTGCTGTTCTTGATCA
ATTTGAGGTATGCTTTGTACATATTCAATTATTTATTTAATTTCCTTTTT
TATTTGCAATATTCTATAAATAATACATTTTATACCCACTATACTAAGAT
AAT.AATTACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGA
AACTCTAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCT
20 TTTGATGGGTAATATAGGCAATTTAAGTTTTATTTCTGTAAAGCAGTAT
TTAGCAAGTACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAT
CTGGTCATTGTACCCAGAATTTAGTTAAATGTAAACATTTTAGATATTAGG
GGACATCAGGTGACAGATATTGTAGAATAGAACAATATATAATATTACCC
AAAACATTTTTTTCTAAGGTTATTCTGTAAATATGTGCTTTCTTGATTT
25 CATTGAATTTGCATTCCTATATTTTAGGTGGTAAAGTGATTGTCTCTTCA
ATAAATCCCGAAATTAATTAAAAAAGAAAAAACAAGTAAATTTTTGA
TATGGAGAGCACTGGTATCATTTAGTATATAAAAAAAGTAAATTTTGAAT
TAAGTTTCTTATATAAAAGCTGTGTATATAGTTTAATTAGTTTACATCA
TTTTTCCATGTGGTGTGTCAGTTGTCTGAAGCAGGTGGTGTCTTGGAG
30 TTTATGCCAATACGCTAGAGAGATARAAATAGKTGGATGCTATGCATTGT
CAAGTGTGATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTG
CTGAGAATAGAGTCTTGTGATGGCATGAAGGAGGTATTTGAACTCAATT
AGGGACGAGCAGCAACAAAAACAACGAGAAGAGTGGTTGCGAGGAAGGAA
TTCCAAGAGTAAATAACAATGTTATTATGCTTCCCAATCTAAAGATATTA
35 AGTATTGGAAATTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACT
TGAAAGCCTGAGACAGCTCCAAGAGTTAAAGATAAAATTTTGCTACGGAA
TGAAAGTGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACA
ACAACAACAACAACGAAGGGGGCATCTTCTTCTTCTTCTTCTTCTTCTC
TTCTTCTTCTAAGAAGGTTGTGGTCTTTCCTTGTCTAAAGTCCATTGTAT
40 TGGTCAATCTACCAGAGCTGGTAGGATTCTTCTTGGGGATGAATGAGTTC
CGGTTGCCTTCATTAGATAAACTTAAGATCAAGAAATGCCCAAAAATGAT
GGTGTTTACAGCTGGTGGGTCCACAGCTCCCCAACTCAAGTATATACACA
CAAGATTAGGCAAACATACTCTTGATCAAGAATCTGGCCTTAACCTTCAT

CAGGTATATATATATTTCTTTAATTGGCATCATCTAATTAAGAAAGATAT
CATTCCTGCCAAGTAAATTTACTTCAAACACATTTCACACTGGTTTCAGTC
TAAGTTTATGTTGTTCTAAGAAGGCCAAAATGGGAAAGCAAGATAGGGAA
AAATAGTGTATTTCAAGTGGAAAGGGTATTTTAGGCATTTTCTGTCAAAAG
5 TTGTTATTGCAGGCTTTTTAGTACCTGGAATCGTGTGTGGGAGGAGCATT
ATTATTCTGATTTGCTTGTTTCTTTATCATTTTTTCTTAGCCTCTCGAAC
AGCTAGAAACCCTTTTAATCTTTTGATTTTCAATGACGAAATTTTTCCCT
GTTACTCCATTTGATTGTTGTTCTTCATGGTTCTAAGTGAGTTATTGGCT
CATCTGTTACTTCTTTTGATTGTTATTTTCATATCATGTTGTCCTTTGAA
10 TCAAGCTTTTCCATTTTCAACCAGGGCAAAAGGTCAAAAGTAACCTACTT
TATGAGATCAAAAACAGCAACCCATCGGATAACTTTTAGTTGGAGTTAAT
AGTTACAATTACCATTGTGATTAATAATTATAATATCTTGTATTAATTCA
TAAAAATTGGTACAGCACATATATGACATTTCAAAGGTTTTTGTTTGACA
TATATATGCCTCTGGCGTTTTCTTTATTGGACTTGCAGACCTCATTCCAA
15 AGTTTATACGGTGACACCTTGGGCCCTGCTACTTCAGAAGGGACAACCTTG
GTCTTTTCATAACTTTATCGAATTAGATGTGGAAGGTAATCATGATGTTA
AAAAGATTATTCCATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAG
ATT.AATGTAAGGTGGTGTAAAAGGGTAGAGGAGGTATTTGAACTGCATT
GGAAGCAGCAGGGAGAAATGGAAATAGTGGAATTGGTTTTGATGAATCGT
20 CAC.AAACAACCTACCACTACTCTTGTCAATCTTCCAAACCTTAGAGAAATG
AACTTATGGGGTCTAGATTGTCTGAGGTATATATGGAAGAGCAATCAGTG
GACAGCATTGAGTTTCCAAACCTAACAAGAGTTGATATCTATAAATGTA
AAAGGTTAGAACATGTATTTACTAGTTCCATGGTTGGTAGTCTATCGCAA
CTCCAAGAGCTACATATATCCAACCTGCAGTGAGATGGAGGAGGTGATTGT
25 TAAGGATGCAGATGATTCTGTAGAAGAAGACAAAGAGAAAGAATCTGATG
GGGAGACGAATAAGGAGATACTTGTGTTACCTCGTCTAAACTCCTTGATA
TTA.AGAGAACTTCCATGTCCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTT
TTC.ATTCCCATTATTGGATACTTTAAGAATTGAGGAATGCCCAGCAATAA
CCACCTTCACCAAGGGAAATTCCGCTACTCCACAGCTAAAAGAAATTGAA
30 ACACATTTTGGCTCGTTTTGTGCTGCAGGGGAAAAAGACATCAACTCTCT
TAT.AAAGATCAAACAACAGGTAAATCAGATCTTTGTTGCTTTAATAATTC
TTAAACTACATTTGAAAAGCTTCATGCAAGTTTTTTTTTGTTATATTGTCA
AAAACCGCAACCTACATTTTCAGCTTTATATTTATGTACTTTATGCAGGA
GTTCAAACAAGACTCTGATTAATGTGAAGTAAATACTAAAGGTAAATTAT
35 ATTTTCATGTTCCCTAGTTGCCTATTAATTAATTGCCTTTTAGTTCATGAT
TTTTGGATGCATTCTTCATGATGATGTCAATCTTCTAATACCCCATTCAT
TGTTTGGTTGAATGTTGACTCTATGTCTTGATGAATATTCAAGGGAAGAA
TTGTTTCATCATATGAAGGACATTAAAGAAGAACATGGATGCTATGAAGAT
GTGGGAAAACAA
40

RG2B deduced polypeptide sequence (SEQ ID NO:90)

MSDPTGLAGAIINPIAQTALVPVTDHVGYMISCRKYVRVMQMKMTELNTSRISVEE
HISRNTRNHLQIPSQTKEWLDQVEGIRANVENFPIDVITCCSLRIRHKLQKAFKITE
QIESLTRQLSLISWTDDPVPLGRVGSMNASTSASLSDDFPSREKTFTQALKALEPNQK
5 FHMVALCGMGGVGKTRMMQRLKKAEEKLFFNYTVGAVIGEKTDPFAIQEAIADY
LGIQLNEKTKPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGLSPFPNQ
GVDFKVLLTSRDSQVCTMMGVEANSIINVGLLTEAEAQSLFQQFVETSEPELQKIGE
DIVRKCCGLPIAKTMACTLRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQ
EEETKSTFLMCGLPEDFDIPTTEELMRYGWGLKLFDRVYTIREARTLRNLCIERLVQ
10 TNLLIESDDVGCVKMHDLVRAFVLGMFSEVEHASIVNHGNMPGWPDENDMIVHSC
KRISLTCKGMIEIPVDLKFPLTILKLMHGDKSLRFPQDFYEGMEKLHVISYDKMKY
PLLPLAPRCSTNIRVLHLTECSLKMFDCCSIGNLSNLEVLSFANSHIEWLPSTVRNLK
KLRLDLRFCDGLRIEQGVLSFKLEEFYIGDASGFIDDNCNEMAERSYNLSALEF
AFFNKAIEVKNMSFENLERFKISVGCSFDENINMSSHSEYENMLQLVTNKGVDVLSK
15 LNGLFLKTEVLFVSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLIISKVELRYLFLK
NLANTLSRLEHLEVCECENMEELIHTGIGGCGEETITFPKLKFLSLSQLPKLSSLCHN
VNIIGLPHLVLDLILKGIPGFTVIYPQNKLRITSSLLKEGVVIPKLETLOIDDMENLEEIW
PCELSGGEKVKLRAIKVSSCDKL VNLFPNPM SLLHHLEELTVENCGSIESLFNIDLD
CVGAIGEEDNKSLLRSINVENLGKLEVVRIKGADNSHLINGFQAVESIKIEKCKRFR
20 NIFTPITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEATGSISNLVFPSCLMH
SFHNLRVLTLDNYEGVEVVFEIESESPTCRELVTTRNNQQQPIILPYLDLYLRNMD
NTSHVWKCSNWNKFFTLPKQQSESPFHNLTINILKCKSIKYLFSPLMAELLSNLKDI
RISECDGIKEVVSNRDDEDEEMTTFTSTHTTTTLFPSLDSLTLFLENLKCIGGGGAK
DEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIEIVGCYALSSVIPCYAA
25 GQMQKL

RG2C polynucleotide sequence (SEQ ID NO:91)

ATAATATTACACAAAGGTAACGTCATTAATTAATTACGATACGAGACAGA
CTTTTTCACTCGGACATNAACGGTCTATTCCTAACTTNANNTAATTNAAT
30 GAATTTAGGATGTGCTAATATGCATGTAANATTCGCTACCGTCATCTTTC
AAATGACCATATTTTTATGTATTTATAATGAATCAATGAAAAACCGGATT
TCTATTTAAAATTCTTAAAACCTTCATCTTTTAAGCCAGGGTGAATACAAT
TGCTAGATCCACTGTTAATTTCCATCGAATTATGCCTGATCAATTGTTGG
CTGCCTACGATGCAGGTGCTACCACAAGAATATGGCCATGGAAACTGCTA
35 ATGAAATTATAAAACAAGTTGTTCCAGTTCTCATGGTTCCTATTAACGAT
TACCTACGCTACCTCGTTTCCTGCAGAAAGTACATCAGTGACATGGATTT
GAAAATGAAGGAATTAAAAGAAGCAAAAGACAATGTTGAAGAGCACAAGA
ATC.ATAACATTAGTAATCGTCTTGAGGTTCCAGCAGCTCAAGTCCAGAGC
TGGTTGGAAGATGTAGAAAAGATCAATGCAAAAGTGGAAGTGTTCCTAA
40 AGATGTCGGCTGTTGCTTCAATCTAAAGATTAGGTACAGGGCCGGAAGGG
ATGCCTTCAATATAATTGAGGAGATCGACAGTGTGATGAGACGACACTCT
CTGATCACTTGGACCGATCATCCCATTCCTTTGGGAAGAGTTGATTCCTG

GATGGCATCCACCTCTACGCTTTCAACTGAACACAATGACTTCCAGTCAA
GAGAGGTAAGGTTTAGTGAAGCACTCAAAGCACTTGAGGCCAACCACATG
ATAGCCTTATGTGGAATGGGGGGAGTGGGGAAGACCCACATGATGCAAAG
GCTGAAGAAGGTTGCCAAAGAAAAGAGGAAGTTTGGTTATATCATCGAGG
5 CGGTTATAGGGGAAATATCGGACCCCATTGCTATTCAGCAAGTTGTAGCA
GATTACCTATGCATAGAACTGAAAGAAAGCGATAAGAAAACAAGAGCTGA
GAAGCTTCGTCAAGGGTTCAAGGCCAAATCAGATGGAGGTAACACTAAGT
TCCTCATAATATTGGATGATGTCTGGCAGTCCGTTGATCTAGAAGATATT
GGTTTAAGCCCTTCTCCCAATCAAGGTGTGCGACTTCAAGGTCTTGTTGAC
10 TTCACGAGACGAACATGTTTGCTCAGTGATGGGGGTTGAAGCTAATTCAA
TTATTAACGTGGGACTTCTAATTGAAGCAGAAGCACAAAGATTGTTCCAG
CAATTTGTAGAACTTCTGAGCCCGAGCTCCACAAGATAGGAGAAGATAT
TGTTAGGAGGTGTTGCGGTCTACCCATTGCCATCAAACCATGGCGTGTA
CTCTAAGAAATAAAAGAAAGGATGCATGGAAGGATGCACTTTCTCGTTTA
15 CAACACCATGACATTGGTAATGTTGCTACTGCAGTTTTTAGAACAGCTA
TGAGAATCTCCCGGACAAGGAGACAAAATCTGTTTTTTTGATGTGTGGTT
TGTTTCCCGAAGACTTCAATATTCCTACCGAGGAGTTGATGAGGTATGGA
TGGGGCTTAAAGTTATTTGATAGAGTTTATACAATTATAGAAGCAAGAAA
CAGGCTCAACACCTGCATTGACCGACTGGTGCAGACAAATTTACTAATTG
20 GAAGTGATAATGGTGTACATGTCAAGATGCATGATCTGGTCCGTGCTTTT
GTTTTGGGTATGTATTCTGAAGTCGAGCAAGCTTCAATTGTCAACCATGG
TAATATGCCTGGGTGGCCTGATGAAAATGATATGATCGTGCCTCTTGCA
AAAGAATTTCAATTAACATGCAAGGGTATGATTGAGTTTCCAGTAGACCTC
AAGTTTCCTAACTAACGATTTTGAACTTATGCATGGAGATAAATCGCT
25 AAAGTTTCCTCAAGAATTTTATGAAGGAATGGAAAAGCTCCGGGTATAT
CATACCATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTCAATGCTCC
ACCAACATTCGGGTGCTTCATCTCACGGAATGTTCAATTAAGATGTTTGA
TTGCTCGTGTATTGGAAATCTATCGAATCTGGAAGTGCTGAGCTTTGCTA
ATTCTTGCAATTGAGTGGTTACCTTCCACGGTCAGAAATTTAAAAAAGCTA
30 AGGTTACTTGATTTGAGATTGTGTTATGGTCTCCGTATAGAACAGGGTGT
CTTGAAAAGTTTGGTCAAACCTGAAGAATTTTATATTGGAAATGCATATG
GGTTTATAGATGATAACTGCAAGGAGATGGCAGAGCGTTCTTACAACCTT
TCTGCATTAGAATTCGCGTTCTTTAATAACAAGGCTGAAGTGAAAAATAT
GTCATTTGAGAATCTTGAACGATTTAAGATCTCAGTGGGATGCTCTTTTG
35 ATGGAAATATCAATATGAGTAGCCACTCATAACGAAAACATGTTGCGATTG
GTGACCAACAAAGGTGATGTATTAGACTCTAACTTAATGGGTTATTTTT
GAAAACAGAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGATCTTGAAG
ATGTTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCATTCTGCAAT
TTAAAAGTCCTTATTATTTCAAAGTGTGTAGAGTTGAGATACCTTTTCAA
40 ACTCAATGTTGCAAACACTTTGTCAAGACTTGAGCATCTAGAAGTTTGTA
AATGCAAGAATATGGAAGAACTCATACTACTGGGATTGGGGGTTGTGGA
GAAGAGACAATTACTTTCCCCAAGCTGAAGTTTTTATCTTTGAGTCAACT
ACCGAAGTTATCAGGTTTGTGCCATAATGTCAACATAATTGGGCTACCAC

ATCTCGTAGACTTGAAACTTAAGGGCATTCCAGGTTTCACAGTCATTTAT
CCGCAGAACAAGTTGCGAACATCTAGTTTGTGGAAGGAAGAGGTAGATAT
ATGTTCTTTATGTTAATAACAATTTAAACAATATTTTCAACCAATTTTCA
TAATATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGGCT
5 ATTAAGGGATGATAATTTTGCAGGTTGTGATTCCTAAGTTGGAGACACTT
CAAATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACTTAG
TGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAAGTGAGTAGCTGTGATA
AGCTTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATCTT
GAAGAGCTTACAGTCGAGAATTGCGGTTCCATTGAGTCGTTATTCAACAT
10 TGACTTGGATTGTGTGCGGTGCAATTGGAGAAGAAGACAACAAGAGCCTCT
TAAGAAGCATCAACGTGGAGAATTTAGGGAAGCTAAGAGAGGTGTGGAGG
ATAAAAGGTGCAGATAACTCTCATCTCATCAATGGTTTTCAAGCTGTTGA
AAGCATAAAGATTGAAAAATGTAAGAGGTTTAGAAATATATTCACACCTA
TCACCGCCAATTTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAGGT
15 TGCGGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTTCATTTCACT
TTCTTAATTAATTANGGACTAAGCTCCTGTTTTTTGAATAATAAAGAGGT
GGGATGACTAACTTGGGCATCACAATTGCAACAAAATGTTACAAACCAT
GAAACGCTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTTAA
AAATATGGCTTAAATTTTTTTTATATTTATGTATCAACATGATTTTTTCATT
20 AGAGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATTAGAACTAA
TTCTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTTAG
TGAAATAAGCAACGGATTTAATAAGTTAACAACCTTAAATGTCATTTCTTA
ACAAAAAAACTATTTGGTTCAGAAAACTGTAATTCAAGATAACTAAAA
TAAAAATATTTGACATTCATAAGAGCATTTTTTTTCTAAATATGATTGCA
25 AATGAATAAAACTTAAATTTATACAGAAAAGATTTTTTATATATGTTATAC
AAAATTTACAAATTGAAATTGGATATGTTAATTAACGGTTTATAATTCTG
GTATCACAAAGGGATATATAATAAAATATTATTTTTCTGTAGTCATTTAT
AATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAGTTA
AACTTTTCATAATAAAAATTTATAATTATTATTTTAAATAAATTATT
30 AATTAAGAGATATATCAAAAATTTAAAGTTATTATAACTTCAAATTTAAC
ATATAATTAATAAATATATGATCATAACTTCCGCAACTCTTCTTTTGTA
TTAAATGACCAGAGAAGCTCTTAGTATATTTTCTAAATCAAAGTCACAA
AACTAATGAAGCATATAATTTTGTGAAAATCAATTAGCATTAGGTTTTAA
GAGTCACCAAAATTCAAAGAGTAATCCAATGCTTTCATTACCACTATGGAG
35 AAAATATTTTCTTAGTTTAAATGAAATGAAAACAAACATTCAAACCTAATT
GTTGCTTATTAAACCAAAGACCCATTACTTAGCCAAGAGTTTAACCAAAA
AAAATTACATTCATGTATCATTATTAATGACTAGATATATATGAATATGA
AGGGAGTTTTTTATAGAAAATATAATCATAGATATTCAACATAACTTCATG
GAATTCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCAACC
40 AAAGAGAAGTTTAGCCTAGCATGGCTAAACTCAAGAAAATAAAAATAAGGA
TTAGAAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTTGTT
CTTGATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAATGCA
AAGGAGAGCCATTAAATTCGTATGTATTGATCCCTTCAAAGCTGCACCA

ACCTCCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCTGAAG
GACCCTATGCGGGTGCCCTGCGCGGGTGGAGCTGAATATGAAAGGTCTTT
GGTCTTTGTGAGGGTGATGTTGTGCGGGTAGCTTGTGCGCATGCTCCGC
GCGGTTGCGGCACATGTGCACAGGTGATGCATGGTGTGTACGTTCTTGAC
5 TTTTGAGCCTCCGATGCTTAGTCCACTTGGCCCAATTCGAGTCCAATCAA
CTTATGACCCATTTTTCTTCAAGTTATCTTCAAGTTAAGCCCAATTTGCC
TTCTCCAAATCATCCATAACTTCACAGAATCGCCCGTTCATCTTAATCCC
GAATGAACAATTATTCTCCCGTCTTCATTTTAAGCAAGATACCACCTTCT
TCATGCTTCATCCATCAATAGTACACTTCATGTATCATCTCTACTAGTTA
10 TTTAGTCCACAGTCCTTGTGTCTCTCAAATTTAATTATCTCATTTAGTT
CCCGTTCCGCTAGTTTCCTTAAAATTTGCAATTAAGCTCACAGAAATATT
AAGTACCCGAAATGGTCATAAAATAACAGAAAGGAAAATATGCATGAAGA
TTAACTAAATGATGAACGAAATATGCTAAAATAGACTATAAAATGAAGTA
AATAAAATGAAATTATCGCACTCCGACCACCTTATAGGCTTGTAGTCCA
15 CCCACCCTTCATTCCTTGTACCAATATGGGATGGAAACATCATTAAATAA
GCCAAAAAATAACATATAAGGGGTGAGTGACAAAGGTAAGTACTAAAGA
TGAAAAAAATCCATTTTTCTTGTATATACACAACACACACATAGGGGCAG
ACGTAGGATTTCATAGTACAGATTGTTGGTGGCACATAAGTGTTGCTAGT
GACATTTTTTTTTTCTTTACGTAGTGGCACAACAGTARAAAAAACRAAA
20 AATTCGAAATTTTTTACAATGTGCCTAAAAAAAACAGTGGTTGTTGGTGC
CACTATGGACACCAAAGTTGAACTGCCCTGCGCGCGCACACACACACAC
ACATAAAGTTTG
GGATGTGATACTTCTTTTGGGAAAATGGAGTTATATCTTTGATATTGTAT
TTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTATAAGTTTTATT
25 TATTTKGATATGAAAAAAAAGTCTTTTATACATTGGATTTAACATAAAA
ATCCAACAATATTAATCAAAAAGACCAACATGTGGACAATTATGTATAT
AATTAATTCACAATAGTCTTTAGGAATAGNATTATATATATAATTAATTC
TCAATGGTCTTAGGAATAGTAAGTTCTTATATTTCAAACNTTTGCCACAN
TTCTTTGNTTACTTNGACACTTTYCTCTMWNNANWMWWTWATATATATAT
30 ATATATATATATAHAHAHAHAVACACACACACTAGATGTGTGCCMGCACA
AAGCAGTGACGTNNNGGAGAANACTTTCTTAAGCATAAATAATTATTATA
TTTTTTATTGGGTATTATATAATAAAAAATTACAACTTTTAAATAAAATA
TTTATGTTTATACTTTATATTTATATTGCTTGTATACTATTAATATAATA
AATTAATATTTATGTCTAATTTATGAAATGTAAATTAATTTAAATACATG
35 AATTTAATATTTTTTAAAATTTTCAGTTTGCTTCAAATTGAGTTTCTTAAT
TATTGACCAAACATGTGGACAATTATGTATATAATTAATTCACAATAGTC
TTTAGGAATAGTATTATATATATAATTAATTCTCAATGGTCTTAGGAATA
GTAAGTTCTTATATTTCAAACCTTTTGCCACAATTCTTTGCTTACTTTGAC
ACTTTTCCTTCCTAACTTTACATATATATATATATTAAGCGCAAAGGTC
40 ATAGGAATATAATATTTTCTATTATTCTACGTTTTGCCACAAAAGTTTGA
ACACTTTGCCACTTTTTGTCCCTCCTTAACCTTTTCAATGTTTTGCGACA
AAAGTTCCAAAACCTTTGCCACTTTGATCATTCCTCAACTTTTCACCGCAT
TAGTTTGTGGAGTTGGCAGTTTTGGTCCCTCTAACTTCGATATTCTCTAC

TGCTAGCCAAAAAGGGTTCAGAGTTTCACACTTTTGGTCCCTGACAGTA
ACCAAATGTGAGATGTCAAATTTTTGCCACATTAGTTTGTGGAGTTGTCC
CTTTTGGTCCCCCACATTCGATATTCTACTATACGATCTTATTTTTCTC
AAATAACAACACGTATATTTAATTACTAATGATAGAAATAGATATCAAAT
5 AAAGTATTTGTAACACTGTGTAGAGTTTTTTTTTACAAGTTTGTATCAAA
TCATATCAAAATTTAAGGTGGAACGGTGACCACATTAACCAGAAATGTAA
TTTATTCTTTGATTTTGATAATTTTTAATATTTTGTGTGATCTATGTAT
TTAAAAGTAAACAACAAGAACATAATCCAAAACCCTAAATTGCAAGTCT
CGCCCAATTTCTCTATCACTAGTCCTCACTTACGATGGCGTTACGTCGCT
10 CTCTCACTGCTTACAACCCTTTGTTGCTACTCATTACAATAACGAAAAGT
TGAATATCCATATATTTATTTGGATGTGGAATTGAACGAATCTCGTCAAA
TTTTTGATTTAGTTGATGGATTTGAGTAGAAGTTTGGGCAGAACGGGAAT
GATGGTCTGCAAGTGGTTATAAACTTGATTCTGAGTTATTACTATATATG
TAGCCTCTTTACAACGACCAAGGTTTCTTCCAGGTACCATTTGATCTTTT
15 TAGAACTTAGTTTTCTGAAACACCCTGATTTGGATCAAATATCACCAACA
ACTCTTAAAAACTTGATTAATCAATTGTTTACTTCATCTTGATAACAAGT
GGAATGATTTTCTACTTGAAAAAAAAGGTCCATGTGCGTCTGGTGGATCT
GGTAAATGAAGATGGAAGGGAGAGCTGACTTTAAAGACACAAACACGTCA
CCATATCTTTTATTTTATTTTAAATTTTCTTTTTTCCTATTTCTTTCTTT
20 CTTGATCTCCAGATGGTATGTGGTGTGGATAATTTACACATAGAGATTGG
GAACGACTGTGATTTAGAGAGGACGTGGCTTGGGGTTGAGGATGGTTTAT
GGCTGGCCGAGTTTCATTTATATAAAACAAACAAATATATAAAACAAGGGG
TAAATGGCCATCTTATATGTATTTAACCGTCCTCTTTTTTTTTTTTTTT
TTTTTTTTTTTTTTTTTTTTTGTAAATTAAGAAGGGGTATACCAGTGTGAGC
25 CTCTTATCCCAACCAGTCAAATAGGGACTTAGGTTGTTTGGAACAGTT
CCGTGAGACCGTGACTTGATGGTAGATAAATTTAGTAACTTAACCCTT
CAATTAACCTACCTTTTTCTTATTAACCTCAATTTCAACCTAAATTCTGAT
TCTTGTGTGAAAATAAGTTGCATCTTTATTTTTGTATTATCTTGTGTCAT
AGGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAAAGATCCAAC
30 ATTTTAAATCTGTTGACGTTTTCCATCATTGCAACTGTTTCTTGAAAAA
AAAATACCTAAAATCAAATAACCATTTTCAAATCCAAAATTATAAGAGA
GAATTGTAAATGGACATGGAATCATAAATCATTAAACACAGTTCAGTAAAC
AAGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAA
AGAGACATTACAAGAAGCCACTGGCAGTATTTCAAATCTTGTATTCCCAT
35 CCTGTCTCATGCACTCTTTTCATAACCTCCGTGTGCTTACATTGGATAAT
TATGAAGGAGTGGAGGTGGTGTGTTGAGATAGAGAGTGAGAGTCCAACAAG
TAGAGAATTGGTAACAACCTCACAATAACCAACAACAGCCTATTATACTTC
CCTACCTCCAGGAATTGTATCTAAGGAATATGGACAACACGAGTCATGTG
TGGAAAGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATC
40 AGAATCACCATTCCACAACCTCACAACCATAGAAATGAGATGGTGTGTCATG
GCTTTAGGTACTTGTTTTCGCCTCTCATGGCAGAAGTTCTTTCCAACCTA
AAGAAAGTCAAGATACTTGGGTGTGATGGTATTAAAGAAGTTGTTTCAAA
CAGAGATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACAAAA

CCACCAACTTGTTCCTCATCTTGATTCTCTCACTCTAAACCAACTGAAG
AATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGA
AATATCTTTCAATAATACCACTGCAACGACTGCTGTTCTTGATCAATTTG
AGGTATGCTTTGTACATATTCAATTATTTATTTAATTTCCTTTTTTATTT
5 GCAATATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAA
TTACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAACTC
TAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGA
TGGGTAATATAGGCAATTTAAGTTTTATTTCTGTAAAGCAGTATTTAGC
AAGTACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGT
10 CATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATTAGGGGTTA
TCAGGTGACAGATATTGTAGAATAGAACAATATGTAATATTACCCAAAAC
TATTTTTTCTAAGGTTGCTCTGTAAATATGTGCTTTCTTGATTTTCATTG
AATTTGCATTCCCTATATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAA
TCCCGAAATTAATTAAAAAAAAAAAAAACAAAAGTAAATTTTTGATATGGA
15 GAGCACTGGTATCATTTAGTATATAAAAAAACTAGATTTTGAATTAAGTT
TCTTATATAAAAGCTGTGTATATAGTTTAATTAGTTTTACATCATTTTTTC
CATGTGGTGTGTCAGTTGTCTGAAGCAGGTGGTGTCTTGGAGCTTATG
CCAATACGCTAGAGAGATAAAAAATAGGCAACTGCCATGCATTGTCAAGTG
TGATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGAGA
20 GTAATGGCTTGCAATGGGATGAAGGAGGTATTTGAAACTCAATTAGGGAC
GAGCAGCAACAAAAACAACGAGAAGAGTGGTTGTGAGGAAGGAATTCCAA
GAGTAAATAACAATGTTATTATGCTTCCCAATCTAAAGATATTAAGTATT
GGA.AATTGTGGGGGTTTGAACATATATTCACATTCTCTGCACTTGAAAG
CCTGAGACAGCTCCAAGAGTTAACGATTAAGGGTTGCTACAGAATGAAAG
25 TGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACAACA
ACAACAACGAAGGGGGCATCTTCTTCTTCTTCTTCTTCTAAGAAGGTGGT
GGTCTTTCCTTGTCTAAAGTCCATTGTATTGGTCAATCTACCAGAGCTGG
TAGGATTCTTCTTGGGGATGAATGAGTTCCGGTTGCCTTCATTAGATAAA
CTTATCATCGAGAAATGCCCAAAAATGATGGTGTTTACAGCTGGTGGGTC
30 CAC.AGCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACTC
TTGATCAAGAATCTGGCCTTAACTTTCATCAGGTACATATATATTCCTTT
AATTGGCATCATCTAATTAAGAAAGATATCATTCCCTGCCAAGTAAATTA
CTTCAAACACATTCACACTAGTTTCAGTCCAAGTTTATGTTGTTCTAGGA
AGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGAGTATTTCAAGTGGA
35 AGGGTATTTTAGGTATTTTCTGTCAAAAATTGTTATTGCAGGCTTTTTAG
TACCTGGAAGAGCATGATTATTCTCGATTGCTTGTCTTTATCATTTTT
TCTTAGCCTAGCATGATTTTCAATGAAATCTTCCCTGTTACTCCATTTG
ATTGTTGTTCTTCATGGTTCTAAGTGAGTTAGTGGCTCATCTGTTACTTC
TTTTGATTGTTATTTTCATAGCATGTTGTCACTTGAATCAAGCTTTTCCA
40 TTTTCAACAAGGACAAAAGGTCAAACTAACCTACTTTATGAGATCAAAA
ATAGCAACCCATCGGATAACTTTTAGTTGGAGTTAATACTTACAATTACC
ATTGTGATTAATAATTATAATATCTTGTATTAATTCATAAAAATTGGTAC
AGC.ACATATATGACATTTCAAAGGTTTTTGTTTGACATATATATGCCTCT

GGCGTTTTCTTTATTGGACATGCAGACTTCATTCCAAAGTTTATACGGTG
ACACCTTGGGCCCTGCTACTTCAGAAGGGACAACCTTGGTCTTTTCATAAC
TTTATCGAATTAGATGTGAAATCTAATCATGATGTTAAAAAGATTATTCC
ATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGTAAAGATTAATGTAATGT
5 GGTGTAAAAGGGTAGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGG
AGAAATGGAAATAGTGGAATTGGTTTTGATGAATCGTCACAAACAACCTAC
CACTACTCTTGTCAATCTTCCAAACCTTGGAGAAATGAAGTTACGGGGTC
TCGATTGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTGAG
TTTCCAAACCTAACAAGAGTTGAAATTTATGAATGTAATTCATTAGAACA
10 TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAG
AGATTGGTTTGTGCAACCATATGGAGGTCGTGCATGTTTCAGGATGCAGAT
GTTTCTGTAGAAGAAGACAAAGAGAAAGAATCTGATGGCAAGATGAATAA
GGAGATACTTGTGTTACCTCATCTAAAGTCATTGAAATTACTACTTCTTC
AAAGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATT
15 TTGGATACTTTGGAAATCTACGAATGCCCAGCAATAACCACCTTCACCAA
GGGAAATTCGCTACTCCACAGCTAAAAGAAATGGAAACAAATTTTGGCT
TCTTTTATGCTGCAGGGGAAAAAGACATCAACTCCTCTATTATAAAGATC
AAACAACAGGTAAACCAGATCTTTGTTGCTTTAATAATTCTTAAACTACA
TTGAAAAGCTTCATGCAAGTTTTTTTTGTTATATTGTCAAAAACCGCAA
20 CCTACATTTTCAGCTTTATATTTATGTACTTTATGCAGGATTTCAAACAA
GACTCTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTTCATGT
TCCTAGTTGCCTATTAATTAAAGGCCTTTTAGTTTCGTGATTTTTGGATGT
ATTCTTCATGATGATGTCAATCTTCTAATACCCCATTCATTGTTTGGTTG
AATGTTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTTCATCA
25 TATGAAGGACATTAAAGAACATGGTGCTAT

RG2C deduced polypeptide sequence (SEQ ID NO:92)

MAMETANEIKQVVPVLMVPINDYLRYLVSCRKYISDMDLKMKEKEDNVEEH
KNHNISNRLEVPAAQVQSWLEDVEKINAKVETVPKDVGCCFNLKIRYAGRDAFNI
30 IEEIDSVMRRHSLITWTDHPIPLGRVDSVMASTSTLSTEHNDFQSREVRFSEALKALE
ANHMIALCGMGGVGKTHMMQRLKKVAKEKRKFGYIIEAVIGEISDPIAIQVVDY
LCIELKESDKKTRAELRQGFKAASDGGNTKFLILDDVWQSVLEDIGLSPSPNQG
VDFKVLLTSRDEHVC SVMGVEANSIINVGLLIEAEAQRLFQQFVETSEPELHKIGEDI
VRRCCGLPIAKTMACTLRNKRKDAWKDALSRHQHDIGNVATAVFRTSYENLPD
35 KETKSVFLMCGLPEDFNIPTEELMRYGWGLKLFDRVYTHIARNRLNTCIDRLVQT
NLLIGSDNGVHVKMHDLVRAFLVGMYSVEQASIVNHGNMPGWPDENDMIVHSC
KRISLTCKGMIEFPVDLKF PKLTILKLMHGD KSLKFPQEFYEGMEKLRVISYHKMKY
PLLPLAPQCSTNIRVLHLTECSLKMFD CSCIGNLSNLEVLSFANSIEWLPSTVRNLK
KLRLLDLRLCYGLRIEQGVLSLVKLEEFYIGNAYGFIDDNCKEMAERSYNLSALEF
40 AFFNNKAEVKNMSFENLERFKISVGCSFDGNINMSSHSYENMLRLVTNKGDVLD SK
LNGLFLKTEVFLSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLISKVELRYLFLK
NVANTLSRLEHLEVCKCKNMEELIHTGIGGCGEETITFPKLKFLSLSQLPKLSGLCH

NVNIIGLPHLVDLKLGIPGFTVIYPQNKLRRTSSLLKEEVVIPKLETQIDDMENLEEI.
WPCELSGGEKVKLREIKVSSCDKLVNLFPRNPMSLLHHLEELTVENC GSIESLFNID
LDCVGAIGEEDNKSLLRSINVENLGKLRVWRIKGADNSHLINGFQAVESIKIEKCK
RFRNIFTPITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEATGSISNLVFPSC
5 LMHSFHNLRVLTLDNYEGVEVVFEIESESPTSRELVTTHNNQQQPILPYLQELYL
NMDNTSHVWKC SNWNKFFTL PKQ QSESPFHNLT TIEMRWCHGFRYLF SPLMAELL
SNLKKVKILGCDGIKEVVSNRDEDEEMTTFTSTHKTTNLFPHLDSLTLNQLKNLK
CIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIKIGNCHAL
SSVIPCYAAGQMQLQVLRVMACNGMKEVFETQLGTSSNKNNEKSGCEE GIPRVN
10 NNVIMLPNLKILSIGNCGGLEHIFTFSALES LRQLQELTIKGCYRMKVIVKKEEDEYG
EQQT TTTTTTKGASSSSSSSKKV VVFPCLKSIVLVNLP ELVGFFLGMNEFRLPSLDKLI
EKCPKMMVFTAGGSTAPQLKYIHTRLGKHTLDQESGLNFHQTSFQSLYGD TLGPAT
SEGTTWSFHNFIELDVKSNDVKKIIPSELLQLQKLVKINVMWCKRVEEVFETALE
AAGRNGNSGIGFDESSQTTTTTLVNL PNLGEMKLRGLDCLRYTWKSNQWTA FEFPN
15 LTRVEIYECNSLEHVFTSSMVGSLLQLQELEIGLCNHMEVVHVQDADV SVEEDKEK
ESDGKMNKEILVPLHLKSLKLLLLQSLKGFSLGKEDFSFPLD TLEIYECPAITTF TK
GNS.ATPQLKEMETNFGFFYAAGEKDINSSI KIKQQDFKQDSD.

RG2D polynucleotide sequence (SEQ ID NO:93) and (SEQ ID NO:94)

20 ACGACCACTATAGGGCGAATTGGGCCCCGACGTCGCATGCTCCCGGCCGCC
ATGGCCGCGGGATGTAAAACGACGGCCAGTCGAATCGTAACCGTTCGTAC
GAGAATCGCTGTCCTCTCCTTCAACCATTTAATGTATATGAGCTAAATTG
AAACATCTACTATCATGTTTAAATTTATAAACTTTTTCTTTAGATTCAC
TTGTCTGGATGTGTTTAATAAAACCCAATTTCCCATATGCGTAGAGATCA
25 TAG.ATGTA ACTATTGTTAATCAATTTTG CCTGCCAAGTTTTAATAATTAT
ACTTG GATATTAACAAAAC TTTATCTAACGACCAAGGTAATATTA AAAAT
AGGTTATTATTCTTCATGCTAATTA AAAGATGGGTTGCAAAAGTGAGACC
ATG.AAAACATTAACACGTTGATATTTTCAACTTTTATTCTTTCATATTCA
CCATATTTTTTTACTTTCTGATTGATTAATCATCTTTCAATCACAGGCTCC
30 TTGGCAAAAAGTCAGATCTATTAACAAATACTTCCATGTGGTTGCAAATT
ACAAGGATTTCAACATAATTACCAAACATAGCATTATCATAAGATCGAA
TAATAATCAAATTCTTCTATAATATTACACAAAGGTAACGTCATTAATTA
ATT.ACGATACGAGACAGACTTTTTCACTCGTGACATCAACGGTCTATTCT
AACTTTACTTAATTAAATGAATCTAGGATGTGCTCATATGCATGTAATAT
35 TTGCTACCGTCATCTTTCAAATGACCATATTTTTATGTATTTATAATGAA
TCAATGAAAAACCGGATTTCTATTTAA AATTCTTAAACTTTCATCTTTTA
AGCCAGGGTGAATACAATTGTAGATCCACTGTTAATTTCCATCGATTATG
CGTGATCAATTGTTGGCTGCATACGATGCAGGTGCTACCACAAGAATATG
GCC.ATGGAACTGCTAATGAAATTATAAAACAAGTTGTTCCAGTTCTCAT
40 GGTTCCCTATTAACGATTACCTACGCTACGTCGTTTCCTGCAGAAAGTACA
TCAGTGACATGGATTTGAAAATGAAGGAATTAA AAGAAGCAAAAGACAAT
GTTGAAGAGCACAAGAATCATAACATTAGTAATCGTCTTGAGGTTCCAGC

AGCTCAAGTCCAGAGCTGGTTGGAAGATGTAGAAAAGATCAATGCAAAAG
TGGAAACTGTTTCCTAAAGATGTCGGCTGTTGCTTCAATCTAAAGATTAGG
TACAGGGCCCGGAAGGGATGCCTTCAATATAATTGAGGAGATCGACAGTGT
CATGAGACGACACTCTCTGATCACTTGGACCGATCATCCCATTCCTTTGG
5 GAAGAGTTGATTCCGTGATGGCATCCACCTCTACGCTTTCAACTGAACAC
AATGACTTCCAGTCAAGAGAGGTAAGGTTTAGTGAAGCACTCAAAGCACT
TGAGGCCAACCACATGATAGCATTATGTGGAATGGGGAGAGTGGGGAAGA
CCCACATGATGCAAAGGCTGAAGAAGGTTGCCAAAGAAAAGAGGAAGTTT
GGTTATATCATCGAGGCAGTTATAGGGGAAATATCGGACCCCATTGCTAT
10 TCAGCAAGTTGTAGCAGATTACCTATGCATAGAGCTGAAAGAAAGCGATA
AGAAAACAAGAGCTGAGAAGCTTCGTCAAGGGTTCAAGGCCAAATCAGAT
GGAGGTAACACTAAGTTCCTCATAATATTGGATGATGTCTGGCAGTCCGT
TGATCTAGAAGATATTGGTTTAAAGCCCTTCTCCCAATCAAGGTGTCGACT
TCAAGGTCTTGTTGACTTCACGAGACGAACATGTTTGCTCAGTGATGGGG
15 GTTGAAGCTAATTCAATTATTAACGTGGGACTTCTAATTGAAGCAGAAGC
ACAAGATTGTTCCAGCAATTTGTAGAACTTCTGAGCCCGAGCTCCACA
AGATAGGAGAAGATATTGTTAGGAGGTGTTGCGGTCTACCCATTGCCATC
AAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGGAAGGA
TGCACCTTCTCGTTTACAACACCATGACATTGGTAATGTTGCTACTGCAG
20 TTTTGTAGAACCAGCTATGAGAATCTCCCGGACAAGGAGACAAAATCTGTT
TTTTTGATGTGTGGTTTGTTCCTCGAAGACTTCAATATTCCTACCGAGGA
GTTGATGAGGTATGGATGGGGCTTAAAGTTATTTGATAGAGTTTATACAA
TTATAGAAGCAAGAAACAGGCTCAACACCTGCATTGAGCGACTGGTGCAG
GCAAATTTACTAATTGGAAGTGATAATGGTGTACACGTCAAGATGCATGA
25 TCTGGTCCGTGCTTTTGTTCCTGGGTATGTATTCTGAAGTCGAGCAAGCTT
CAATTGTCAACCATGGTAATATGCCTGGGTGGCCTGATGAAAATGATATG
ATCGTGCACTCTTGCAAAAGAATTCATTAAACATGCAAGGGTATGATTGA
GATTCCAGTAGACCTCAAGTTTCCTAAACTAACGATTTTGAAACTTATGC
ATGGAGATAAGTCTCTAAAGTTTCCTCAAGAATTTTATGAAGGAATGGAA
30 AAGCTCCAGGTTATATCATACGATAAAATGAAGTACCCATTGCTTCCTTT
GGCACCTCAATGCTCCACCAACATTTCGGGTGCTTCATCTCACTGAATGTT
CATTAAAGATGTTTGATTGCTCTTCTATCGGAAATCTATCGAATCTGGAA
GTGCTGAGCTTTGCTAATTCTCGCATTGAATGGTTACCTTCCACAGTCAG
AAATTTAAAGAAGCTAAGGTTACTTGATCTGAGATTTTGTGATGGTCTCC
35 GTATAGAACAGGGTGTCTTGAAAAGTTTGGTCAAACCTGAAGAATTTTAT
ATTGGAAATGCATATGGGTTTATAGATGATAACTGCAAGGACATGGCAGA
GCGTTCTTACAACCTTTCTGCATTAGAATTCGCGTTCTTTAATAACAAGG
CTGAAGTGAAAAATATGTCATTTGAGAATCTTGAACGATTCAAGATCTCA
GTGGGGTGCTCTTTTGATGGAAATATCAGTATGAGTAGCCACTCATACGA
40 AAACATGTTGCAATTGGTGACCAACAAAGGTGATGTATTAGACTCTAAAC
TTAATGGGTTATTTTGAAGGAGGTTGCTTTTTTTAAGTGTGCATGGC
ATGAATGATCTTGAAGATGTTGAGGTGAAGTCGACACATCCTACTCAGTC
CTCTTCATTCTGCAATTTAAAAGTCCGTATTATTTCAAAGTGTGTAGAGT

TGAGATACCTTTTCAAACCTCCATGTTGCAAACACTTTGTCAAGCCTTGAG
CATCTAGAAGTTTGTGGATGCGAAAATATGGAAGAACTCATACATACTGG
GATTGGGGGTTGTGGAGAAGAGACAATTACTTTCCCAAGCTGAAGTCTT
TATCTTTGAGTCAACTACCGAAGTTATCAGGTTTGTGCCATAATGTCAAC
5 ATAATTGGGCTACCACATCTCGTAGACTTGAACTTAAGGGCATTCCAGG
TTTCACAGTCATTTATCCGCAGAACAAGTTGCGAACATCTAGTTTGTGTA
AGGAAGAGGTAGATATATGTTCTTTATGTTAATAACAATTTAAATAATATT
TTCAACCAAATTTTCATAATATATCTGTAATTTGATTGTATGATGTGTTA
TTGTTTATATGTGGCTATTAAGGGATGATTATTTTGCAGGTTGTGATTCC
10 TAAGTTGGAGACACTTCAAATTGATGGCATGGAGAACTTAGAAGAAATAT
GGCCTTGTGAGCTTAGTGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAA
GTGAGTAGCTGTGATAAGCTTGTGAATCTATTTCCGCACAATCCCATGTC
TCTGTTGCATCATCTTGAAGAGCTTAAAGTCAAAAATTGTCGTTCCATTG
AGTCGTTATTCAACATCGACTTGGATTGTGTCAGTGCAATTGGAGAAGAA
15 GAC.AACAAGAGCATCTTAAGAAGAATCAAAGTGAAGAATTTAGGGAAGCT
AAGAGAGGTGTGGAGGATAAAAGGTGCAGATAACTCTCGTCCCCTCATCC
ATGGCTTTCCAGCTGTTGAAAGCATAAGTATCTGGGGATGTAAGCGGTTT
AGAAATATATTCACACCTATCACCGCCAATTTTGATCTGGTGGCACTTTT
GGAGATTCACATAGGAAATTACAGAGAAAATCATGAATCGGAAGAGCAGG
20 TAACGCTTTCAATTTCACTTTCTTACTTAATTAAGGACTAAGCTCTTGTT
TTTTGAATAATAAAGAGGTGGGATGACTAACTTGGGCATCACAATTGTA
ACAAAATGTTACAAACCATGAACGTACAAACCATTTCTTGAATTAAGGTT
TCAATACAAGTCATTTACAAATATGGCTTAAGTTTTTTTATATTTATGTA
TCAACATTATTTTTCATTAGAGGTCATTATTATAATAGTAAGTTTAAAGC
25 AATTTAAATTAGCACTAATTTTTTCATCATCTAACTTTAGCTAATAAATCG
TTATAAATGTCAATAGCTAAAATAAAAATATTTGACATTCCTGAGAGCA
ATTTTTTCTAAACATGATTGCAAATGATTAAACTTAAATTTAAACTAAA
AAGATTTTTATATATGTTATACAAAATTTACAAATTGAAATTGGATATGT
TAATTAACAGTTTATAATTATTGTATTACAAAGCGATATATAATAAATA
30 TTATTTTTCTGTAGTCATGTATAATTGTATATGTAAATGATTTTTTAAGA
TGGTAGAAGTGGAACCTAGTCAATCTCACTTAACTCATTGTACACCAGT
TTTATATCCGTTTCTCTCTCTCTCTCTTCTTGCCTCCATCTTTTTTCAAC
TCATAACACATAAAAATAACATATTTTCCAACACATTTAAGTCACTACCA
CATCATTATTTTTTAATTTAATTAAATTAGAAAATATAAAATTAAATAAAA
35 CAT.AACATTTTTTTATTTAAAAGGCACTAATACAAATAAAAAGATACACGG
TAAATAAAAAAACGATAATTAGAAAAAAACATAATAAAAAAAGACAACA
TTA.AAAATAWAAAGCGACAATAAAATTAATAATGATCAAGAAAATTCT
AAAACCTCCACCATATTTTTCTGCAATTTGTCATTTATGTTCAAACACCA
TTCGCAGAATCCCTCCTATCAAGTGATCATGTTGATTGAGAAAAAACTGT
40 ATGTCTCTCATGTATCTCCAAGTCCAACAAGTTAGCTTTTCAATTTCTT
ATTTTCTCATGTAAGACGCAAATTTTCATCCCGATATTGTTTTCTATCTT
CCACCTCTACTTTATTCACAGTGTGGATGAAGGAGAGGACAGCGATTCTC
GTACGAACGGTTACGATTCGACTGGCCGTCGTTTTACAATCCCGCGGCCA

TGGCGGCCGGGAGCATGCGACGTCGGGCCCATTGCGCCTATAGTGGTCGT
AATACA (SEQ ID NO:93)

Sequence gap

5 TGAGCCTCCGATGCTTAGTCCACTTGGCACAGTTCAAGTCCAATCAACTT
ATAACCCATTTTTCTTCAAGTTGTCTTCAAGTTAAGCCCAATTTGCCTTC
TCCAAATCATCCATAACTTCATGGAATCGCCCCCTTCATCTTAATCCCGAA
TGCACAATTATTCTCCCATCTTCATTTTAAGCAAGAGGGCCACCTTCTTCA
TGCTTCATCCATCAATAGTCTGTTGGAATAGTGTCTAAGGCTGCAACTAT
ATTAGACAAGTATTTGACCCGGTTGTGCATGGTCCTTTTGGGTTGCCTTC
10 ACCATAGCAACTTGATAGGATGATTTATTAAGAGAGAGTAAATATTATTA
ATATATTATGAGAATAATATAATGAATAATATATTTGTTATTTGATTAAT
ATAAGTCATAGAATTAATTAGAATTAATTTGGTGACTTAAAGAGATTAAT
TAAATAAAGGGGTATAAACTGTCAATTGTTTGATAGTTAAGCTTTAGACT
GTAAATCCATTTGGATATGGTATGGACGAATCCTAAGGGATTTAGGATAG
15 CTAAAATCGTCCATATGAGTTATCTAAGAAGGATTTGGATAGCCTTAAGA
GAAGATTATCTGATAGGGACTTATCTGTAATCCTTAAGGAGTCTACAAGT
ATAAATAGACCCTATGGCTGATGGAATTCGACACATCTCCTAAAGTAAGA
GAGCCTTGGCCGAATTCCTCCCCTCACCTCTCTCCTAAATCATTCTTCTT
GCTATTGGTGTTTGTAAGCCATTAGAGGAGTGACATTTGTGACTCTAGAA
20 TCTCCAAGACCTCAAGATCAACAAGGAATTCAAAGGTATGATTCTAGATC
TGTTTCAATGTTGTTATTTGTCCTAATTAGTCATTAGAAGACTTGGATTCT
AAAGCATGTTTATTAGAAAGCCTAGATCYGAGCAATAGGGTTTTGCATGC
GCACATAGGAAAGTTCTTATGGCTAAAACCCATCATAGTCCACTTCATGT
ATCATCTCTACTAGTTATTTAGTCCATAATCCTTGTTGTCCTCCAAGTTT
25 AATTACCTCCCTTAGTTCCCTGTTCTGCTAGTTTCCTTAAAATTTGCTATT
AAGATCACAGAAGTACAGAGTACCCAAAATGGTTATAAAATAACAAAAAG
GAAAATATGCATGAAGATTAATAATTATAAATGTAATATGCTAAAATA
AACTATAAAAAAAAAGTAAATAAAATGAACTATCACACTCCGACCACCC
TTATAGGCTTGTACTGCACCCACCCTTCATTCCCTTGTAACCAATATGGGAT
30 GGAACATTATTCATTAAGCCAAAAAACTAACATTTAAGGGGTGAGTGAC
AAAGGTAAGTACTAAAGACAACAATAATCCATTTTTCTTGTAACATACACA
ACACACACATAGGGGCGGACGTAGGATTTGTAGTATGTGTTGTGGGTGAC
ACATTTTTTTCTTTTACGTAGTGACACAATAGTAGAGAAAACGAGAAATTC
CAATTTTTTTACATTGTGTTTCGAAAAAATATACAGGGGTGCTGGTGCTAC
35 TCTGGGCACCAAAGTGGAACCGCCCCTGCACACACACACACATAGAGGGA
GAGAGAGAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT
TGGGATGTGATACTTCTTTTGGGAAAATGGAGCAATATCTTTAATATTGT
ATTTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTATAACTTTTA
GTTTTTTTTTATTTAATCTGTATATTTAATCATTTTCAAGTTTATAAGTTTT
40 ATTTATTTTGGTATACCAGAAAAAAAAGTCTTTTATGTGTTGGATTTAAC
ATAAAAATCTAACAATATTAATCAAAAAGACCAACATGTGGACAATTAT
GTATATAATTAATTCTCAATGGTCTTAGTGTAACGATATAAATTTCAAAA
CAATTTTTTCACATTAAAAAAAACACTTTCAGTCATAATTGTTATAAATTA

TCATTGTATCACAAAATCAGTTCATAACATCACATCCCAAGATCAATAAA
GTGTAAATACTCCTCATGTGTGTACTAATCAAGCCGACGCCTTCCCGCGA
TTCTCACTGGTACCTGAAACACGTAACATAACAACCTGTAAGCATAAATGC
TTAGTGAGTTCCCCAAAATACCACATACCACATATATGCCTTTCCAGGCC
5 ATA.ACTCTGTAGGATCTTCCGACCCAAGTGTCTCAGGGGACTTCCGTCCC
GAATCCCGGTAGACCTTCCGGTCCTACCCGTATTGACCTTCCGGTCCGTA
TCATACATAACATAACATAACACATACATATCACATAACAACATATAGCAC
ATACATCTCATAACATAAAAGACCTTCCGGTCACATAAAGGTACCCTTCC
AGGTACAGTATAGTGAGAACTCACCTCGTATGATGTCTAATACCTCAC
10 GTGCTCGATATCCCTGAATCTCGAAACAATGACCTAGCCCCGCCTACTCA
CATAAAGTAATTATTTCAAATCATTAACGGCTCTCAAGGCTAGACTACAT
CCCTTTCTATAAATCCACAGAAGGGTAAAAGACCATTTTACCCCTCCTTG
ACCCAAAAGTCCAAATGTTGATCAAAACCCCAAAAGTCAACGAAAGACAA
TGGTCAACTTTGACCCTACTCGTGGAGTGCACAAAGGTGACTCGGCAAGT
15 ACATGCGGGTCCTCTGAATCCTTTCAGTCTCTCTTGGCTCGTCGAGTCTT
TCTTCCACCCGACGAGTTACACCTGTCATGAATCGCGGGGCAACCCCGAC
TCG.ACTTGTGAGTCCGCTCATGGACTCAACGAGTTCATTCCATGCTCAC
ACTCAAATGACCTCCTGAGGTCAGATCTGTTCTCTAATCCATAGATCTG
ACCTTCCCAAGCTCAATAAACACGTAAAGGTTCGAACTTGATACTCATGC
20 AACGTCCAAATGATTCTACTTGATGATTTAGCCCCAAATACAACATCCTA
AGTCCATACGACCTTATTTTTCTCAAATAACAACACATATATTTAATTAC
CAATGACAGTAATAGATATCATATAAAGTATTTGTAACACTTTGTAAGAA
CCTTGCTACTATAGGTAAAAAGAAACATTTCAAAGTACATGCCCTAATTA
GAAAAAAAGTTATAAAAAAATAATGACTAGGGGCGTGTTTTTTTACTAG
25 TTTGTATCAAATTATATCAAATTTAAGGTGGAAAAGAATGACGACCACA
TTA.ACCAGAAATGTAATTATTTTTTTATTTGGTAATTTTAAATATTTGTT
GTG.ATCTATGTATTTAAAAGTAAATATCAAACAAGAACATAATCCAAACC
CTA.AATTGCAAGTCTCGCCCAATTTCTCTATCACTAGTCCTCACTTACGA
TGGCGTTACGTCGCTCTCTCACTTCCTACAACCCATTGTTGCTACTAATT
30 A.CACTAACGAAAAGTTGAATATCCATATATTTATTTGGATGTGAAATTGA
ACG.AATCTCGTCAAATTTTTTATTTTGTGATGGATTTGAGTGGAAGTTT
AGGCAGAACGGGAATGATGGTCTGCAAGTGGTTATAAACATGGGTGAAGA
TAA.AATGGAGTTGTCGCCGTTGTATTATAGATCTCTTAGGGGTTTGATT
TGAGTTATTACTGTATACGTAGCCTCTTTACAACGACCATTCTTCCAAGT
35 ACCATTTGATCTTTTTTAGAATCCAGTTGTCTGAAACACCCTGATTTGGAT
CAAATATCACCAACAACCTCTTAAGAACTGGACTAATTAATTGTTTTCTTG
ATCTTGATAACAAGAGGAAACACGTCACCATATCTTTTATTTTAAATTTG
CTTTTGGTGTATTTTCTTTCTTCCCATTTCTTTCTTGATCTGTTCCAGAT
GGTATTTGGTGTGGATAATTTACACCTGGAGATTGTGAACGATGGGAAGG
40 GGTATGTGATTTACAGAGGATGTGGCTTGTGGTTGAGGATGGTTTATGGC
TGGCCGAGTCTAATTTATATTTATATAAACAATAAATATATAAAACAAG
GGTAAAATATGTATTTAAGCGTCCTCTTTAATGGTGACAATTTTACAG
TTT.ACTCTCTTTGTTTTTTAATTGTGATGCCACGATCGAACTCATTCAT

CCCCCCCCCTTTTTTTTTTAAAATAAAAAATTAAGAAGGGGTACCACCAT
ATACCCGTGTCAGCTTCTTATTCCCAAGCAGTCAAATAGGGACTTAGGTT
GTATGGAAACAGTTCCGTGACTTGGATGGCAGATAAATTTAGTAAACTTA
ACCCTTCAATTAACCTACCTTTTTCTTATTAACCTCAATTTCAAGCTAAAT
5 TCTGATTCTTGTTTGAAAATAAGTTGCATCTTTATTTTTGCATATTATCT
TGTTGCATAGGATCCTTAGCATCTTTTAATAGTTTATTTGAAGCTGAAAG
ATCCAACCTAGTTTTGATCTGTTGGCATTTTCCATCATTTGCAACTGTTTC
TTGAAAAAAAATACCTAAAATCAAAATAACCATTTTCAAATCCAAAATTA
TAAGAGAGAATTGTTAATGGACGTGGAATCATAAATCATTAACACAGTTC
10 AGTACACAAGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTAT
CAGAGAAAGAGACATTACAAGAAGTCACTGATACTAATATTTCTAATGAT
GTTGTATTATTCCCATCCTGTCTCATGCACTCTTTTCATAACCTCCATAA
ACTTAAATTGGAAAATTATGAAGGAGTGGAGGTGGTGTGTTGAGATAGAGA
GTGAGAGTCCAACATGTAGAGAATTGGTAACAACCTACAATAACCAACAA
15 CAGCCTATTATACTTCCCAACCTCCAGGAATTGTATCTAAGGAATATGGA
CAACACGAGTCATGTGTGGAAGTGCAGCAACTGGAATAAATTCTTCACTC
TTCCAAAACAACAATCAGAATCACCATTCCACAACCTCACAACCATAGAA
ATGAGATGGTGTGCATGGCTTTAGGTACTTGTTTTCGCCTCTCATGGCAGA
ACTTCTTTCCAACCTAAAGAAAGTCAAGATACTTGGGTGTGATGGTATTG
20 AAGAAGTTGTTTCAAACAGAGATGATGAGGATGAAGAAATGACTACATTT
ACATCTACCCACACAACCACCAACTTGTTCCCTCATCTTGATTCTCTCAC
TCTAAAATACATGCACTGTCTGAAGTGTATTGGTGGAGGTGGTGCCAAGG
ATGAGGGGAGCAATGAAATATCTTTCAATAATACCACTACAACCTACCGAT
CAATTTAAGGTATGTTTGTACATATTTAATTATATATTTAATTTCTTGT
25 TAATTTCTTTTCTTTGCAATATTCTATGCGAACTCAAGAATGGGATTG
GAGGCATATAAAGTTACATTCATTTGAACAAGTATTACCTTTTATTTGTT
ATTTATCATTTTCATATCAAGTACCTATAACATTTCTTTTTTATTTTTCT
AATTAGAAGAGGTCCACATGTCTAATTAGGTTTTCCATTCTATGTGTAAC
CTCTATTCTCTCTGTAATCAAGCATCTTAGATTATTTATCCATTTTCATA
30 ATTGTTGTTTATTTTTACAGTTTTTTTTTTTTATTTAATTTTAATAATTTAA
TTTTAATTTATTTATTTTTTTTTTTTTGGTAATTGCAACCTGTCATATAT
TCAAGTCTTAATGTAACATAATAACATTTTATACCCACTATACTAAGA
TAATAATTACCTAAAGGGATGGATGCCATGACACTGCTACACTTCAGNAA
CTCTAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTT
35 TGATGGGTAAATATAGGCAATTTAAGTTTTATTTCTGTAAAGCAGTATTT
AGCTAGTAGTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTCAAATCT
GGTCATTGTACCCAGAATTTAGTTAAATGTAACATTTTAGATATTAGGGG
TCATCAGGTGACAGATATTGTAGAATAGAACAATATGTAATATTACCCAA
AACTATTTTTTCTAAGGTTGCTCTGTAAATATGTGCTTTCTTGATTTC
40 TTGAATTTGCATTCGTATATTTTAGGTGGTAAACTGATTGTCTCTTCAAT
AAATCCTGAAATTAATTAACAAAAAACAACAAAGTACATTTTGTATT
GGAGAGCACTGGTATCATTTAGTATAGAAAAAACTAGATTTTGAATTAY
CTTTCTTATATAAAAGTTGTGTATATAGTTTAATTAGTTTTACATCATTT

TTCTATGTGTTGTTGCAGTTGTCTGAAGCAGGTGGTGTGTTTGTGAGCTT
ATGCCAATACTCTAGAGAGATAGAGATATATAGGTGTGATGCACTGTCAA
GTGTAATTCCATGTTACGCAGCAGGACAAATGCAAAAGCTGCAAGTGCTG
ACAGTCAGTTCTTGTAATGGTCTGAAGGAGGTATTTGAAACTCAATTAGG
5 GACGAGCAGCAACAAAAACAACGAGAAGAGTGGTTGTGAGGAAGGAATTC
CAAGAGTAAATAACAATGTTATTATGCTTCCCAATCTAAAGATATTGGAA
ATCTACGGTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGA
AAGCCTGAGACAGCTCCAAGAGTTAACGATTAAGGGTTACTACTCTTGTC
AATCTTCCAAACCTCAAAGAAATGAGGTGGAGTGGCTAAGTAATCTGAG
10 GTATATATGGAAGAGCAATCAGTGGACAGCATTTGAGTTTCCAAACCTAA
CAAGAGTTGAAATTTGTGAATGTAATTCATTAGAACATGTATTTACTAGT
TCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTACATATATTTAACTG
CAGTCTGATGGAGGAGGTAATTGTTAAGGATGCAGATGTTTCTGTAGAAG
AAGACAAAGAGAAAGAATCTGATGGCAAGACGAATAAGGAGATACTTGTC
15 TTACCTCATCTAAAGTCCTTGAAATTACAACCTCTTCGAAGTCTTAAGGG
GTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTATTGGATACTTTAG
AAATCAAAAGATGCCCAACAATAACCACCTTCACCAAAGGAAATTCGCT
ACTCCACAATAAAAGAAATACAAACAATTTTGGCTTCTTTTATGCTGC
AGGGGAAAAAGACATCAACTCTCTTATAAAGATCAAACAACAGGTAAATC
20 AGATCTTTGTTGCTTTAATAATTCTTAAACTACATTTGAAAAGCTTCATG
CAAGTTTTTTTTGTTATATTGTCAAAAACCGCAACCTACATTCAGCTTTAT
ATTTATGTACTTTATGCAGGATTTCAAACAAGACTCAGATTAATGTGAAG
TGAATATTAAAGGTAAATTATATTTTCATGTTCCCTAGTTGCCTATTAATT
AATGGCCTTTTAGTTCATGATTTTTTGGATGTATTCTTCATGATGATGTGA
25 ATCTTCTAATACCCCATTCATTGTTTGGTTGAATGTTGACTCTATGTCAG
GATGAATATTCAAGGGAAGAATTGTTTCATCAWATGAAGGACATTAAAGAA
CATGGATGCTATGAAGATGTTGGGAAAACATATGTATCAAGTGGCAARCT
GCTTAATGATCTAAGTTTGTGGTTGANGATGTTGATTTTAATATTTCAA
ATTCATTGGTTATATGGGCTTATCAATAGTGTTAATGGGATAATGAGTGA
30 CTTAACCTAAATTATGTTGTTGGTAAATGTTGGACAAGTATGGAAAATTA
GGAATGACTTGTGAAAAAAAATAAAAAAAA (SEQ ID NO:94)

RG2D deduced polypeptide sequence (SEQ ID NO:95)

MAMETANEIKQVVPVLMVPINDYLRYVVSCRKYISDMDLKMKEKAKDNVEE
35 HKNHNISNRLEVPAAQVQSWLEDVEKINAKVETVPKDVGCCFNLKIRYRAGRDAF
NIIIEIDSVMRRLITWTDHPIPLGRVDSVMASTSTLSTEHNDFQSREVRFSEALKA
LEANHMIALCGMGRVVGKTHMMQRLKKVAKEKRKFGYIEAVIGEISDPIAIQQVVA
DYLCELKESDKKTRAEKLRQGFKAKSDGGNTKFLIILDDVWQSVDLedigLSPSPN
QGVDFKVLLTSRDEHVC SVMGVEANSIINVGLLIEAEAQR LFQQFVETSEPELHKIG
40 EDIVRRCCGLPIAKTMACTLRNKRKDAWKDAL SRLQHHDIGNVATAVFRTSYENL
PDKETKSVFLMCGLFPEDFNIPTEELMRYGWGLKLFDRVYTIIEARNRLNTCIERLV
QANLLIGSDNGVHV KMHDLVRAFVLGMYSEVEQASIVNHGNMPGWPDENDMIVH

SCKRISLTCKGMIEIPVDLKF PKLTILKLMHGD KSLKFPQEFYEGMEKLQVISYDKM
KYPLLPLAPQCSTNIRVLHLTECSLKMFDCCSIGNLSNLEVLSFANSRIEWLPSTVRN
LKKLRLLDLRFCDGLRIEQGVLSVKLEEFYIGNAYGFI DDNCKDMAERSYNLSA
LEFAFFNKA EVKNMSFENLERFKISVGCSFDGNISMSSH SYENMLQLVTNKGDVL
5 DSKLNGFLFKTEVLF LSVHGMNDLEDVEVKSTHPTQSSSFCNLKVRIISKCVELRYL
FKLHVANTLSSLEHLEVCGCENMEELIHTGIGGCCGEETITFPKLKSLSLSQLPKLSGL
CHNVNIIGLPHLVDLKLKGIPGFTVIYPQNKLR TSSLLKEEVVIPKLET LQIDGMENL
EEIWPCELSGG EKVKLREIKVSSCDKL VNLFPHPNMSLLHHLEELKVKNCRSIESLF
NIDLDCVSAIGEEDNKSILRRIKVKNLGKLREVWRIKGADNSRPLIHGFPAVESISIW
10 GCKRFRNIFT PITANFDLVALLEIHIGNYRENHESEEQIEILSEKETLQEVTD TNISND
VVLFPSCLMHSFHN LHKLKLENYEGVEVVFEIESESPTCRELVTTHNNQQQPILPN
LQELYL RNMDNTSHVWKCSNWNKFFTL PKQQSESPFHNLT TIEMRWCHGFRYLFS
PLMAELLSNLKKVKILGCDGIEEVVSNR DDEDEEMTTFTSTHTTTNLFPHLDSLTLK
YMHCLKCIGGGGAKDEGSNEISFNNTTTT TDQFKLSEAGGVCWSLCQYSREIEIYRC
15 DALSSVIPCYAAGQMQLQVLT VSSCNGLKEVFETQLGTSSNKNNEKSGCEE GIPR
VNNNVIMLPNLKILEIYGCGGLEHIFTFSAL ESLRQLQELTIKGY YTLVNLPNLKEM
RLEWLSNLRYIWKSNQWTA FEFPNLTRVEICECNSLEHVFTSSMVGSL LQLQELHIF
NCSLMEEVIVKDADV SVEEDKEKESDGKT NKEILVLP HLKSLKLQLLRS LKGFSLGK
EDFSFPLD TLEIKRCPTTTTFTK GNSATPQLKEIQTNFGFFYA AGEKDINSLIKIKQQ
20 DFKQDSD.CEVNIK

RG2E polynucleotide sequence (SEQ ID NO:96)

TGGGAAGACACAATGATGCAAAGGTTGAAGAAGGTTGCTAAAGAAAATAGAAT
GTTCAATTATATGGTTGAGGCAGTTATAGGGGAAAAGACAGACCCACTTGCTAT
25 TCAACAAGCTGTAGCGGATTACCTTTGTATAGAGTTAAAAGAAAGCACTAAACC
AGCAAGAGCTGATAAGCTTCGTGAATGGTTTAAGGCCAACTCTGGAGAAGGTA
AGAATAAGTTCCTTGTAATATTTGATGATGTTTGGCAGTCCGTTGATCTGGAAG
ACATTGGTTTAAGTCATTTTCCAAATCAAGGTGTCGACTTCAAGGTCTTGTTGA
CTTCACGAGACGAACATGTTTGCACAGTAATGGGGGTTGAAGCTAATTCAATTC
30 TTAATGTGGGACTTCTAGTAGAAGCAGAAGCACAAAGTTTGTTCAGCAATTTG
TAGAAACTTTTGAGCCCGAGCTCCATAAGATAGGAGAAGATATCGTAAGGAAG
TGTTGTGGTTTACCTATTGCCATTA AAAACCATGGCATGTACTCTAAGAAATAAA
AGAAAGGATGCATGGAAGGATGCACTTTTGCATTTAGAGTACCATGACATTAGC
AGTGTTGCGCCCAAAGTCTTTGAAACGAGCTACCATAATCTCCACAACAAGGAG
35 ACTAAATCTGTGTTTTTGTATGTGTGGTTTTTTTCTGAAGACTTCAATATTCCAA
TCGAGGAGTTGATGAGGTATGGATGGGGCTTAAAGATATTTGATAGAGTTTATA
CTATTAGACAAGCAAGAATCAGGCTCAACACCTGCATTGAGCGACTGGTG CAG
ACAAATTTGTTAATAGAAAGTGATGATGGTGTGCACGTCAAGATGCATGATCTG
GTCCGTGCTTTCGTTTTGGTTATGTTTTCTGAAGTTGAACATGCTTCAATTATCA
40 ACCATGGTAATATGCTTGGATGGCCTGAAAATTATATGACCAACTCTTGCAAAA
CAATTTCAATTAACATGCAAGAGTATGTCTGAATTTCCGGGAGATCTCAAGTTTC
CAAACCTAACGATTTTGAAACTCATGCATGGAGATAAGTTGCTAAGATATCCTC

AAGACTTTTATGAAGGAATGGAAAAGCTCTGGGTTATATCATATGATGAAATGA
AGTATCCATTGCTTCCCTCGTTACCTCAATGCTCCATCAACCTTCGAGTGCTTCA
CCTCCATCGATGCTCATTAAATGATGTTTGATTGCTCTTGTATTGGAAATATGTTG
AATCTGGAAGTGCTTAGCTTTGTTAAATCTGGCATTGAATGGTTACCTTCCACA
5 ATAGGAAATTTAAAGAAGCTAAGGTTACTTGATCTGAGAGATTGTTATGGTCTT
CGTATAGAAAAAGGTGTCTTGAAAAATTTGGTGAAAATTGGAGGAATTTATATT
GGTAGAGCAGATATTTTATAGAT

RG2E deduced polypeptide sequence (SEQ ID NO:97)

10 WEDTMMQRLKKVAKENRMFNYMVEAVIGEKTDP LAIQQAVADYLCIELKESTKP
ARADKLRWFKANS GEGKNKFLVIFDDVWQSV DLEDIGLSHFPNQGVDFKVL LTS
RDEHVCTVMGVEANSILNVGLLVEAEAQSLFQQFVETFEPELHKIGEDIVRKCCGL
PIAKTMACTLRNKRKDAWKDALLHLEYHDISSVAPKVFETSYHNLHNKETKSVFL
MCGFFPEDFNIPIEELMRYGWGLKIFDRVYTIRQARIRLNTCIERLVQTNLLIESDDG
15 VHVKMHDLVRAFVLVMFSEVEHASIINHGNMLGWPENYMTNSCKTISLTCKSMSE
FPGDLKFPNLTILKLMHGDKLLRYPQDFYEGMEKLWVISYDEMKEYPLLPSPQCSI
NLRVLHLHRCSLMMFDCSCIGNMLNLEVL SFVKS GIEWLPSTIGNLKKLRLLDLRD
CYGLRIEKGVLKNLVKIGGIYIGRADIL.

20 RG2F polynucleotide sequence (SEQ ID NO:98)

CTGTGGAAGACACAATGATGCAAAGGCTGAAAAAGGTTGTGCATGAAAAGAAA
ATGTTTAACTTTATTGTTGAAGCAGTTATAGGGGAAAAGACAGACCCCGTTGCC
ATT CAGGATGCTATAGCAGATTACCTAGGTGTAGAGCTCAATGAAAAATCTAAG
CAAGCAAGAGCTGATAAGCTCCGTCAAGGATTCAAGGACAAATCAGATGGAGG
25 CAAAAATAAGTTCCTTTGTAATACTTGACGATGTTTGGCAGTCTGTTGATCTGGA
AGATATTGGTTTAAAGTCCTTTTCCAAATCAAGGCGTCGACTTCAAGGTCTTGTT
GACATCACGAGACAGACATGTTTGCACAGTGATGGGGGTTGAAGCCAAATTAA
TTCTAAACGTGGGACTTCTAATTGAAGCTGAAGCACAAAGTTTGTTCACCAAT
TTGTTGTCACCTCTGAGCCCGAGCTCCATAAGATAGGAGAAGATATTGTAAAGA
30 AGTGTTTCGGTCTGCCAATTGCCATCAAAACCATGGCATGTACTCTACGACATA
AAAGAAAGGATGCATGGAAGGATGCACTTTCACGTTTAGAGCACCATGACATT
CAAAGTGTTGTGCCTAAAGTATTTGAAACGAGCTACAACAATCTCAAAGACAA
GGAGACTAAATCCGTATTTTGTATGTGTGGTTTGTTCCTGAAGACTTGGATAT
ACCTATCGAGGAGTTGATGAGGTATGGATGGGGCTTAAGATTATTTGATAGAGT
35 TAATACTATTACACAAGCAAGAAACAGGCTCAACACCTGCATTGAGCGACTGG
TGCACACAAATTTGTTAATTGAAAGTGTTGATGGTGTGCATGTCAAGATGCATG
ATCTGGTTCGTGCTTTTGT TTTGGGAATGTTTTCTGAAGTGGAGCATGCTTCAAT
TGTCAACCATGGTAATATGCCCAGGTGGACTGAAAATGATATGACTGACTCTTG
CAAACAAATTTTATTAACATGCAAGAGTATGTTGGAGTTTCCTGGAGACCTCAA
40 GTTTCCAAACCTAAAGATTTTGAACCTTATGCATGGAGGTAAGTCACTAAGGTA
TCCTCAAGACTTTTATCAAGGAATGGAAAAGCTGGAGGTTATATCATACGATGA
AATGAAGTATCCATTGCTTCCCTCGTTGCCTCAATGTTCCACCATCCTTCGAGTG

CTTCATCTCCATGAATGTTTATTAAGGATGTTTGATTGCTCTTCAATCGGTAATC.
TTTTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAGCATTGAATTGTTACCTTC
CGTAATTGGAAATTTGAAGAAGTTGCGGCTGCTAGATTTGACAAACTGTTATGG
TGTTTCGTATAGAAAAGGATGTCTTGAAAAATTTGGTGAAACTTGAAGAGCTTTA
5 TATTAGGAATGGTCTACCAGTTTACAGAGGAT

RG2F deduced polypeptide sequence (SEQ ID NO:99)

VEDTMMQRLKKVVHEKKMFNFIVEAVIGEKTDPVAIQDAIADYLGVELNEKSKQA
RADKLRQGFKDKSDGGKNKFFVILDDVWQSVLEDIGLSPFPNQGVDFKVLTSRD
10 RHVCTVMGVEAKLILNVGLLIEAEAQSLFHQFVVTSEPELHKIGEDIVKKCFGLPIAI
KTMACTLRHKRKDAWKDALSRLEHHDIQSVVPKVFETSYNNLKDKETKSVFLMCG
LPEDLDIPIEELMRYGWGLRFLDRVNTITQARNRLNTCIERLVHTNLLIESVDGVH
VKMHDLVRAFVLGMFSEVEHASIVNHGNMPEWTENDMTDSCQISLTCKSMLEFP
GDLKFPNLKILKLMHGGKSLRYPQDFYQGMKLEVISYDEMKYPLLPSLPQCSTILR
15 VLHLHECSLRMFDCSSIGNLNFNMEVLSFANSSIELLPSVIGNLKKLRLDLTNCYGV
RIEKDVLKNLVKLEELYIRNGLPVYRG

RG2G polynucleotide sequence (SEQ ID NO:100)

GAAGACACGATGATGAAGAACTGAAGGAGGTCGTGGGACAAAAGAAATCATTC
20 AATATTATTATTCAAGTGGTCATAGGAGAGAAGACAAACCCTATTGCAATTCAG
CAAGCTGTAGCAGATTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGC
AAGAGCTGATAAGCTTCGTAAACGGTTTGAAGCCGATGGAGGAAAGAATAAGT
TCCTTGTAATACTTGACGATGTATGGCAGTTTGTCTGATCTTGAAGATATTGGTTT
AAGTCCTCTGCCAAATAAAGGTGTCAACTTCAAGGTCTTGTTGACGTCAAGAGA
25 TTCACATGTTTGCACCTCTGATGGGAGCTGAAGCAAATTCAATTCTTAATATAAA
AGTTTTTAAAGATGTAGAAGGACAAAGTTTGTTCGCCAGTTTGCTAAAAATGC
GGGTGATGATGACCTGGATCCTGCTTTCAATGGGATAGCAGATAGTATTGCAAG
TAGATGTCAAGGTTTGCCCATGCCATCAAACCATTCCTTAAGTCTTAAAGG
TAGAAGCAAGTCTGCATGGGACGTTGCACTTTCTCGTCTGGAGAATCATAAGAT
30 TGGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAATTAGCTACGACAATCTCCA
AGATGAGGTTACTAAATCTATTTTTTTACTTTGTGCTTTATTTCTGAAGATTTT
GATATTCCTACTGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATA
GAAGCAAAAACCTATAAGAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCG
GCTTAGGGAGACAAATTTGTTATTTGGAAGTGATGACATTGGATGTGTCAAGAT
35 GCACGATGTGGTGCGTGATTTTGTGTTTGCATATATTCTCAGAAGTCCAACACGC
TTCAATTGTCAACCATGGTAACGTGTCAGAGTGGCTAGAGGAAAATCATAGCAT
CTACTCTTGTAAGAAATTTCAATTAACATGCAAGGGTATGTCTCAGTTTCCCAA
AGACCTCAAATTTCCAAACCTTTCAATTTTGAAACTTATGCATGGAGATAAGTC
ACTGAGCTTTCTGAAAACCTTTTATGGAAAGATGGAAAAGGTTTCAAGGTAATATC
40 ATATGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACCAA
CGTTCGAGTGCTTCATCTTCATTACTGTTTCAATTAAGGATGTTTGATTGCTCTTCA
ATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAACATTGAA

TGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACA
AATTGTAAAGGTCTTCGTATAGATAATGGTGTCTTAAAAAATTTGGTCAAACCTT
GAAGAGCTTTATATGGGTGTTAATCGTCCGTATGGACAGGCCGTTAGCTTGACA
GATGAAAA

5

RG2G deduced polypeptide sequence (SEQ ID NO:101)

RHDDEELKEVVGQKKSFNIIIQVVIGEKTNP IAIQQA VADYLSIELKENTKEARADKL
RKRFEADGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTSRDSHVCTL
MGAEANSILNIKVLKDVEGQSLFRQFAKNAGDDDLDPAFNGIADSIASRCQGLPIAI
10 KTIALSLKGRSKSAWDVALSRLENHKIGSEEVVREVFKISYDNLQDEVTKSIFLLCAL
FPEDFDIPTEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVK
MHDVVRDFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSQFPKDL
KFPNLSILKLMHGDKSLSPENFYGKMEKVQVISYDKLMYPLLPSSLECSTNVRVLH
LHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRID
15 NGVLKNLVKLEELYMGVNRPYGQAVSLTDE

RG2H polynucleotide sequence (SEQ ID NO:102)

TGAAGGAGGTTGTGGAACGAAAGAAAATGTTCAGTATTATTGTTCAAGTG
GTCATAGGAGAGAAGACAAACCCTATTGCTATTCAGCAAGCTGTAGCAGA
20 TTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATA
AGCTTCGTAAATGGTTCGAGGCCGATGGAGGAAAGAATAAGTTCCTTGTA
ATACTTGACGATGTATGGCAGTTTGTGCGATCTTGAAGATATTGGTTTAAG
TCCTCTGCCAAATAAAGGTGTCAACTTCAAGGTCTTGTGACGTCAAGAG
ATTCACATGTTTGCACCTCTGATGGGAGCCGAAGCCAATTCAATTCTCAAT
25 ATAAAAGTTTTAACAGCTGTAGAAGGACAAAGTTTGTTCGCCAGTTTGC
TAAAATGCGGGTGATGATGACCTGGATCCTGCTTTCAATAGGATAGCAG
ATAGTATTGCAAGTAGATGTCAAGGTTTGCCCATTGCCATCAAAACCATT
GCCTTAAGTCTTAAAGGTAGAAGCAAGCCTGCGTGGGACCATGCGCTTTC
TCGTTTGGAGAACCATAAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTT
30 TTA-AAATTAGCTATGACAATCTCCAAGATGAGATTACTAAATCTATTTTT
TTACTTTGTGCTTTATTTCTGAAGATTTTGATATTCCTACTGAGGAGTT
GATGAGGTATGGATGGGGCTTGAAATTATTTATAGAAGCAAAAACCTATAA
GAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCGGCTTAGGGAGACA
AATTTGTTATTTGGAAGCGATGACATTGGATGCGTCAAGATGCACGATGT
35 GGTGCGTGATTTTGTGTTTGCATATATTCTCAGAAGTCCAGCACGCTTCAA
TTGTCAACCATGGTAACGTGTCAGAGTGGCTAGAGGAAAATCATAGCATC
TACTCTTGTAAGAATTTCAATTAACATGCAAGGGTATGTCTGAGTTTCC
CAAAGACCTCAAATTTCCAAACCTTTCAATTTGAAACTTATGCATGGAG
ATAAGTCGCTGAGCTTTCCTGAAAACCTTTATGGAAAGATGGAAAAGGTT
40 CAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCCCTCATCACT
TGAATGCTCCACTAACGTTTCGAGTGCTTCATCTCCATTATTGTTCAATAA
GGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTC

AGCTTTGCTAATTCTAACATTGAATGGTTACCATCTACAATTGGAAATTT
GAAGAAGCTAAGGCTACTAGATTTGACAAATTGTAAAGGTCTTCGTATAG
ATAATGGTGTCTTAAAAAATTTGGTCAAACCTTGAAGAGCTTTATATGGGT
GTTAATCATCCGTATGGAC

5

RG2H deduced polypeptide sequence (SEQ ID NO:103)

KEVVERKKMFSIIVQVVIGEKTNPQIAIQQAVADYLSIELKENTKEARADKLRKWFEA
DGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTSRDSHVCTLMGAEAN
SILNIKVLTAVEGQSLFRQFAKNAGDDDLDPAFNRIADSIASRCQGLPIAKTIALSLK
10 GRSKPAWDHALSRLNHNKIGSEEVVREVFVKISYDNLQDEITKSIFLLCALFPEDFDIP
TEELMRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVVR
DFVLHIFSEVQHASIVNHGNVSEWLEENHSIYCKRISLTCKGMSEFPKDLKFPNLSI
LKLMMHGDKSLSPENFYGKMEKVQVISYDKLMYPLLPSSLECSTNVRVLHLHYCSL
RMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKGLRIDNGVLKN
15 LVKLEELYMGVNHYPYG

RG2I polynucleotide sequence (SEQ ID NO:104)

AAGAAGAGCTGAAGGAGGTTGTGGAACAAAAGAAAACGTTCAATATTATT
GTTCAAGTGGTCATAGGAGAGAAGACAAACCCTATTGCTATTCAGCAAGC
20 TGTAGCAGATTCCCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAA
GAGCTGATAAGCTTCGTAAATGGTTCGAGGCTGATGGAGGAAAGAATAAG
TTCCTCGTNATACTTGACGATGTATGGCNGTTTGTGATCTTGAAGATAT
TGGTTTAAAGTCCTCATCCAAATAAAGGTGTCANCTTCAAGGTCTTGTTGA
CGTCAAGAGATTCACATGTTTGCACCTCTGATGGGAGCTGAAGCCAATTCA
25 ATTCTCAATATAAAAGTTTTAAAGATGTAGAAGGAAAAAGTTTGTTCGG
CCAGTTTGCTAAAAATGCGGGTGATGATGACCTGGATCCTGCTTTCATTG
GGATAGCAGATAGTATTGCAAGTAGATGTCAAGGTTTGCCCATTCGCATC
AAAACCATTCGCTTAAGTCTTAAAGGTAGAAGCAAGTCTGCATGGGACGT
TGCACCTTCTCGTCTGGAGAATCATAAGATTGGTAGTGAAGAAGTTGTGC
30 GTGAAGTTTTTAAATTAGCTATGACAATCTCCAAGATGAGGTTACTAAA
TCT.ATTTTTTTACTTTGTGCTTTATTTCTGAAGATTTTGATATTCCTAC
TGAGGAGTTGGTGAGGTATGGGTGGGGCTTGAAATTATTTATAGAAGCAA
AAACTATAAGAGAAGCAAGAAACAGGCTCAACACCTGCACTGAGCGGCTT
AGGGAGACAAATTTGTTATTTGGAAGTGATGACATTGGATGCGTCAAGAT
35 GCACGATGTGGTGCGTGATTTTGTTCATATATTCTCAGAAGTCCAGC
ACGCTTCAATTGTCAACCATGGTAATGTGTCAGAGTGGCTAGAGGAAAAT
CATAGCATCTACTCTTGTAAGAATTTTCAATTAACATGCAAGGGTATGTC
TGAGTTTCCCAAAGACCTCAAATTTCCAAACCTTCAATTTTGAACTTA
TGCATGGAGATAAGTCGCTGAGCTTTCCTGAAAACCTTTATGGAAAGATG
40 GAAAAGGTTTCAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCC
CTCATCACTTGAATGCTCCACCAACCTTCGAGTGCTTCATCTCCATGAAT
GTTCAATTAAGGATGTTTGATTGCTCTTCAATTGGTAATCTTCTCAACATG

GAAGTGCTCAGCTTTGCTAATTCTGGCATTGAATGGTTACCATCTACAAT
TGGAAATTTGAAGAAGCTAAGGCTACTGGATCTGACAGATTGTGGAGGTC
TTCATATAGATAATGGCGTCTTAAAAAATTTGGTCAAACCTTGAAGAGCTT
TATATGGGTGCTAATCGTCTGTTTGGAAAGTGCCAT

5

RG2I deduced polypeptide sequence (SEQ ID NO:105)

EELKEVVEQKKTFNIIVQVVIGEKTNP IAIQQA VADSLSIELKENTKEARADKLRKWF
EADGGKNKFLVILDDVW?FVDLEDIGLSHPNKG V?FKVLLTSRDSHVCTLMGAEA
NSIL.NIKVLKDVEGKSLFRQFAKNAGDDDLDPAFIGIADSIASRCQGLPIAKTIALSL
10 KGRSKSAWDVALSRLENHKIGSEEVREVFKISYDNLQDEVTKSIFLLCALFPEDFDI
PTEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVV
RDFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEFPKDLKFPNLS
ILKLMHGDKSLSFPENFYGKMEKVQVISYDKLMYPLLPSSLECSTNLRVLHLHECSL
RMFDCSSIGNLLNMEVLSFANSIEWLPSTIGNLKKLRLLDLTDCGGLHIDNGVLKN
15 LVKLEELYMGANRLFGKCH

RG2J polynucleotide sequence (SEQ ID NO:106) and (SEQ ID NO:107)

ATGTCCGACCCAACAGGGATTGTTGGTGCCATTATTAACCCAATTGCTCA
AACGGCCTTGGTTCCCCTTACAGACCATGTAGGCTACATGATTTCTGCA
20 GAAATATGTGAGGGACATGCAAATGAAAATGACAGAGTTAAATACCTCA
AGAATCAGTGCAGAGGAACACATTAGCCGGAACACAAGAAATCATCTTCA
GATTCCATCTCAAATTAAGGATTGGTTGGACCAAGTAGAAGGGATCAGAG
CGAATGTTGCAAACCTTTCCAATTGATGTCATCAGTTGTTGTAGTCTCAGG
ATCAGGCACAAGCTTGGACAGAAAGCCTTCAAGATAACTGAGCAGATCGA
25 AAGTCTAACGAGACAAAATTCGCTGATTATCTGGACTGATGAACCTGTTT
CCCTGGGAAGAGTTGGTTCCATGATTGCATCCACCTCTGCAGCATCAAGT
GATCATCATGATGTCTTCCCTTCAAGAGAGCAAATTTTTAGGAAAGCACT
AGAAGCACTTGAACCCGTCCAAAAATCCACATAATAGCCTTATGGGGGA
TGGGCGGAGTGGGGAAGACCACGATGATGAAGAAGCTGAAAGAGGTCGTG
30 GAACAAAAGAAAACGTGCAATATTATTGTTCAAGTGGTCATAGGAGAGAA
GACAAACCCTATTGCTATCCAGCAAGCTGTAGCAGATTACCTCTCTATAG
AGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTAAACGG
TTCGAAGCCGATGGAGGAAAGAATAAGTTCCTTGTAATACTTGACGATGT
ATGGCAGTTTTTTCGATCTTGAAGATATTGGTTTAAGTCCTCTGCCAAATA
35 AAGGTGTCAACTTCAAGGTCTTGTTGACGTCAAGAGATTCACATGTTTGC
ACTCTGATGGGAGCTGAAGCCAATTCTATTCTCAATATAAAAGTTTTAAA
AGATGTAGAAGGAAAAAGTTTGTTCGCCAGTTTGCTAAAAATGCGGGTG
ATGATGACCTGGATCCTGCTTTCATTGGGATAGCAGATAGTATTGCAAGT
AGATGTCAAGGTTTGCCCATTGCCATCAAAACCATTCCTTAAGTCTTAA
40 AGGTAGAAGCAAGTCTGCATGGGACGTCGCACTTCTCGTCTGGAGAATC
ATAAGATTGGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAAATTAGCTAT
GACAATCTCCAAGATGAGGTTACTAAATCTATTTTTTTACTCTGTGCTTT

ATTTCTGAAGATTTTGATATTCCTATTGAGGAGTTGGTGAGGTATGGGT
GGGGCTTGAAATTATTTATAGAAGCAAAAACATAAGAGAAGCAAGAAAC
AGGCTCAACAACTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGG
AAGTCATGACTTTGGGTGCGTCAAGATGCACGATGTGGTGCGTGATTTTG
5 TTTTGCATATGTTTTTCAGAAGTCAAGCATGCTTCAATTGTCAACCATGGT
AACATGTCAGAGTGGCCAGAGAAAAATGATACCAGCAACTCTTGTA AAAAG
AATTTCAATTAACATGCAAGGGTATGTCTAAGTTTCCTAAAGACATCAACT
ATCCAAACCTTTTGATTTTGAACTTATGCATGGAGATAAGTCGCTGTGC
TTTCTGAAAACTTTTATGGAAAGATGGAAAAGGTTTCAGGTAATATCATA
10 TGATAAATTGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACTA
ACGTTTCGAGTGCTTCATCTCCATTATTGTTTCAATTAAGGATGTTTGATTGC
TCTTCAATTGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTC
TAACATTGAATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGC
TACTAGATTTGACAAATTGTAAAGGTCTTCGTATAGATAATGGTGTCTTA
15 AAAAATTTGGTCAAACCTTGAAGAGCTTTATATGGGTGTTAATCGTCCGTA
TGGACAGGCCGTTAGCTTGACAGATGAAAACCTGCAATGAAATGGTAGAAG
GTTCCAAAAA ACTTCTTGCACTAGAATATGAGTTGTTTAAATACAATGCT
CAAGTGAAGAATATATCCTTCGAGAATCTTAAACGATTCAAGATCTCAGT
GGGATGTTCTTTACATGGATCTTTCAGTAAAAGCAGGCACTCATACGAAA
20 ACACGTTGAAGTTGGCCATTGACAAAGGCGAACTATTGGAATCCCGAATG
AACGGGTTGTTTGAGAAAACGGAGGTTCTTTGTTTAAGTGTGGGGGATAT
GTATCATCTTTCAGATGTTAAGGTGAAGTCCTCTTCGTTCTACAATTTAA
GAGTCCTTGTCGTTTCAGAGTGTGCAGAGTTGAAACACCTCTTCACACTT
GGTGTGCAAATACTTTGTCAAAGCTTGAGCATCTTAAAGTCTACAAATG
25 CGATAATATGGAAGAACTCATAACACCGGGGGTAGTGAAGGAGATACAA
TTACATTCCCAAGCTGAAGCTTTTATATTTGCATGGGCTGCCAAACCTA
TTGGGTTTGTGTCTTAATGTCAACGCAATTGAGCTACCAAACTTGTGCA
AATGAAGCTTTACAGCATTCCGGGTTTCACAAGCATTATCCGCGGAACA
AGTTGGAAGCATCTAGTTTGTGAAAGAAGAGGTACATATACATATAGTT
30 TATGTTAATACATTTTAAACAATCTTTTCAACTAAAAGTTTCAGAATATA
TCTGTATTTTGATTGTATGATGTGTTAGTGTTTGGATGTGGCTATTAAAG
GATAATTATTTGGCAGGTTGTGATTCCTAAGTTGGATATACTTGAAATTC
ATGACATGGAGAATTTAAAGGAAATATGGCCTAGTGAGCTTAGTAGAGGT
GAGAAAGTTAAGTTGAGAAAGATTAAAGTGAGAAATTGTGATAAACTTGT
35 GAATCTATTTCCACACAATCCCATGTCTCTGCTGCATCATCTTGAAGAGC
TTATAGTCGAGAAATGTGGTTCCATTGAAGAGTTGTTCAACATCGACTTG
GATTGTGCCAGTGTAATTGGAGAAGAAGACAACAACAGCAGCTTAAGAAA
CATCAATGTGGAGAATTCAATGAAGCTAAGAGAGGTGTGGAGGATAAAAAG
GTGCAGATAACTCTCGTCCCCTCTTTCGTGGCTTTCAAGTTGTTGAAAAG
40 ATAATCATTACGAGATGTAAGAGGTTTACAAATGTATTACACCTATCAC
CACAAATTTTGATCTGGGGGCACTTTTGGAGATTTTCAGTTGATTGTAGAG
GAAATGATGAATCAGACCAAAGTAACCAAGAGCAAGAGCAGGTATGGATT
TCAATTTTACTCTTTTACTTAATTAATGATTAAGCCCCTGCTTTTTTAATA

AAAAGGGGACAAACCATTTCTTGACTTAATGTTGCAATACAAGTCATGTA
TAAGAGTGATTAACCTTTTTTTTATTTATAAAATAACTACAAAACATGTTT
TTTCATTATAGATCATGTATAAATGTGACTAATTTTTTTCATCGCCTAAC
TTTTGTTGATAAATCATTAGAAATGTCACTAATTACTTTTTAGTATTTAT
5 AAAATAACTACAAAACATGTTTTTTCATTATAGATCATGTATATATCAAC
TAAAAATATTATTCCCTTACACAAAAAAGGTTCAAGAAAGCCTGTA
TTTCGAAATAACTAAAAAGAAAATATTTGATATTCACTAAGAGAAATTTT
TTTCTAAACATGATCGCAAATGATTAAACTTAAATTAAACTAAAAAGA
TTTTTATATATGTTATNCAAAATTAATAATTTGAAATTAAGTTTATAATTC
10 TNGTNTCACAAAGGGATATATATAGTAAAATATTATTTTTTGCAGTCAT
GCATAGTTGTATTTTTTAAATGATTTATTAACGTGGTAGGAGTGGAACCA
CTCAATCTAGTAGACCCACTATCACATGTCACATCAGCTTTACATCTATT
TTTCTTTCTCCTTTTTTTCATCTTTTTTAACTCATAACACNTAAAANTANC
ATATTTTCCAACACACTNAACTCATTGTCACATTATTATTTTTTAATTTAA
15 TTAATTNGAAAATTAATAATTAANTAAANCNTAACATTTTTTAATTAATA
AATATTAATCCAAATAAAAAANTNCACGATAAATTAATAAANGTTTANTTTG
GAAAAAANCC (SEQ ID NO:106)

Sequence gap

ATAACCTTTCAAGGGTCAACTCAAGTCCAAGTTAAAGTCAAGGTCAAAA
20 CCTTGGTTAAAGTCAACTTTGGTCAAAGTCAACATCTACTTGACTCACCT
CACCGAGTTGGTCCACCAACTTGTCGAGTCCCTTAATCCACAACTTCAA
GAACCTTCGATCCTACTCGTCGAGTCTTTCAAGAACTCTTCGAGTTCCAT
TACACAGAATCGGGACCTTTTGCTCATGACTCGCCGAGTTCATCCTTGAA
CTTGTCGAGTCTAGCTTCATACGAGTTCGAGTGTTTAGTCCTTGACTCGT
25 CGAGTTCTTCCTTGAACCTCGTCGAGTCCATCTTCGTATAGTTGGGACATT
GCCTTGAACCTCACCGAGTTCATCATTGAACTCATCGAGTCCTTCGATCTT
CAAGTCCATAATCCTGTCCATCTTGTTGAGTCCTCTTCTAGACTCAACCA
GATTCCTCAGAAACAGAAAAGGTTAGGGAACCATTACCTGACTCGCCGAG
TCCCAAGAACGAATCCCCGAGTCCCCCAATGTCCATGACCATACAATCGA
30 TTTTCGTTGGGCTCATTGCATCCAAAGCATAGATCTAACCTCCTAGGGTC
CATATTACACGTAAAGCTACGAACTTGACGTCCATGCATGGGGGATTTGG
CTC.AAATGGCATTAAAATGGGGTTTATCTGATGCATGGGACTCCCATGGC
CATAAAGTTAACACCTTTATGCCATGGGAATCCTCAATGGTTCCATATCT
GAAGTTAACACTCTACAATATGTTCTAAACCCGAAGGTGGCTTAGAAATG
35 CCCCAAAATGGCAAGATTCAAGCCTTAAAGGAGATCTAACAAATGATAAG
TCAAGGTTCAAGCTTTTTACCTTGAATAAGCTGGAAATGAAGCAAAATCT
CTGGATCCACTTGCTTCTTCAAGAACCCCCAAGCTTCCACTTCTTCCTTC
AAGTTTCAAACAACTTTAAACACTCAAAAATGGCTCAAGAACACTCAAAA
AGCTTTAGGGTTTCGAGTTAGGGCTTTTTGGAAGCGAGAGGGACGATGGG
40 GGCTGAAATGAGGCTAGAAAAAGTGTTTAAATAGGGGGCAAACCCTAAAT
ATTAGGGTTTCATCCAGGCAGCCCTACTCGTCGAGTCGGGCTCCCGACTC
GTCGAGTAGGTCACTTAAAACCCGCGTCCATAATCCAGTCTACTCGACGA
GTTGGGCTCCAACCTCGTCGATTCCGAGTGCAAAACGTTCAATTACTTAA

ATTTAAATATGTACCAGGAACCGGGTGTACAGTTGAGACTTTATACCTC
CATAAGATAGATCTAGGTGCACATAGCCTGGATCCACAAGCTCCATGTCA
ACAAGCGACTCTTCAAGAAGTTCATTCTTCCTCCTTAAGCACCAAAAAAC
ACACAAAATCACCATGAAGCTCAAGAAATACTCAAATAGAGGATAGGGTT
5 TCGTTCGTAGGGTTAGAGAGGATGGAGGCTAGAGGAAATGAGGGATAGAG
GCGAGTTAAGGTCTTTAAATAGGGTCCAAGACCCTAAATTAGGGTTTTAA
TCTGGCCAGACGAACGCAGGGTGTTCCTCAAATGCATATGTGTCCAAATTC
TCGTGTGCGCCATGCGTACCTCCCTTGTACGCCATGTGTACCGGGTTTGG
TCCAAACCCTTCTAACTTCAAATGATCATAACTTGCACCCCTTATCTGTT
10 TTCGATGTTCTTTATATCCACGGAAAGGTAACAAGAAGCCCTATACTTCT
ATAAACTTTATTTAATCTGAAAACCAACCGAAATTAAATCCAAAATTCAT
AAAAGTCCCGAACCAACACATTTACCGATACCCTTGGGCTCCAAAACACA
AATTGAAAACCCGGATCATCCAACTACATCATCCACCTCCAAATGAGCC
CAAACCTCAATTATTCAAGGGTTCTAAGCCTGTTAATGCCCACTCCTCGAT
15 TACCACCCCGCAATGGGAAACGATTCAAAACAGGGCGTTACATAATTTGT
TGTGGTTTTGTATTTTTTATTTCCGGTGAAGGTGAAAGATCCAACATTTT
TTAATCTGTTGGCATTTTCCATCATTTGCAACTGTTTCTTGAAAAA
TACCTAAAATCAAATAACCATTTTCAAATCCAAAATTATAAGAGAGAAT
TGTAATGGACATGGAATCTTAAATCATTAAACACAGTTCAGTACACAAGT
20 TGCTAATTACATTTCTTGCTGTGCAGATTGAAATTCTATCAGAGAAAGAG
ACATTACAAGAAGCCACTGACAGTATTTCTAATGTTGTATTCCCATCCTG
TCTCATGCACTCTTTTCATAACCTCCAGAACTTATATTGAACAGAGTTA
AAGGAGTGGAGGTGGTGTGAGATAGAGAGTGAGAGTCCAACAAGTAGA
GAATTGGTAACAACCTCACCATAACCAACAACAGCCTGTTATATTTCCCAA
25 CCTCCAGCATTTGGATCTAAGGGGTATGGACAACATGATTCGCGTGTGGA
AGTGCAGCAACTGGAATAAATTCTTCACTCTTCCAAAACAACAATCAGAA
TCCCCATTCCACAACCTCACAACCATAAATATTGATTTTTGCAGAAGCAT
TAAGTACTTGTTTTACCTCTCATGGCAGAACTTCTTTCCAACCTAAAGA
AAGTCAATATAAAATGGTGTATGGTATTGAAGAAGTTGTTTCAAACAGA
30 GATGATGAGGATGAAGAAATGACTACATTTACATCTACCCACACAACCAC
CATCTTGTTCCCTCATCTTGATTCTCTCACTCTAAGTTTCCTGGAGAATC
TGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAAATA
TCTTTCAATAATACCACTGCAACTACTGCTGTTCTTGATCAATTTGAGGT
ATGCTTTGTTTCATATTCAATTATTTATTTAATTTCTTTTTTATTTGCAA
35 TATTCTATAAATAATACATTTTATACCCACTATACTAAGATAATAATTAC
CTAGAGGGGATGGATGCTATGACACAGCTGCTACACTTCAGAACTCTAGT
AAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTTTGATGGG
TAATATAGGCAATTTAAGTTTTATTTCTGTAAAGCAGTATTTAGCAAGT
ACTGGCCAGTAGGAGAGGAGAATATCACCTTTTGTGAAAATCTGGTCATT
40 GTACCCAGAATTTAGTTAAATGTAACATTTTAGATATCAGGGGACATCAG
GTGACAGATATTGTAGAATAGAACAATATATAATATTACCCAAAACCTATT
TTTTCTAAGGTTTTTCTGTAAATATGTGCTTTCTTGATTTCAATTGAATT
TGCATTCCTATATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAATCCC

GAAATTAATTAAAAACAAAAGTAAATTTTGGATATGGAGA
GCACTGGTATCATTTAGTATATAAAAAAAGTATTTTGAATTAAGTTTC
TTATATAAAAGCTGTGTATATAGTTTAATTAGTTTTACATCATTTTTCCA
TGTGGTGTTCAGTTGTCTGAAGCAGGTGGTGTTCCTTGGAGCTTATGCC
5 AATACGCTAGAGAGATAAGTATAGAATTCTGCAATGCATTGTCAAGTGTG
ATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTTCAAGTGCTGACAGT
CAGTTCTTGTAAATGGTCTGAAGGAGGTATTTGAAACTCAATTAAGGAGGA
GCAGCAACAAAAACAACGAGAAGAGTGGTTGTGATGAAGGAAATGGTGGA
ATTCCAAGAGTAAATAACAATGTTATTATGCTTTCTGGTCTGAAGATATT
10 GGAAATCAGCTTTTGTGGGGGTTTGGAACATATATTCACATTCTCTGCAC
TTGAAAGCCTGAGACAGCTCGAAGAGTTAACGATAATGAATTGCTGGTCA
ATGAAAGTGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAAC
AACAAACAACGAAGGGGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
CTTCTTCTTCTTCTTCTTCTCCTCCTTCTTCTTCTAAGAAGGTTGTGGTC
15 TTTCTTGTCTAAAGTCCATTGTATTGGTCAATCTACCAGAGCTGGTAGG
ATTCTTCTTGGGGATGAATGAGTTCGGTTGCCTTCATTAGATGAACCTTA
TCATCGAGAAATGCCCAAAAATGATGGTGTTCACAGCTGGTGGGTCCACA
GCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACTATTGA
TCAAGAATCTGGCCTTAACCTTTCATCAGGTATATATGTTTCTTTAATTGG
20 CATCATCTAATTAAGAAAGATATCATTCTGCCAAGTAAATTTACTTCAA
ACACATTCACACTGGTTTCAGTCTAAGTTTATGTTGTTCTAGGAAGGCCA
AAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTTCAAGTGGAAAGGGTA
TTTTAGGTATTTTCTGTCAAAAGTTGTTATTGCAGGCTTTTTAGTACCTG
GAATCGTGTGTGGGAGGAGCATTATTATTCTGATTTGCTTGTTCCTTTAT
25 CATTTTTCTTAGCCTCTCGAACAGCTAGAAACCCTTTAATCTTTTGAT
TTTAAATGACAAAATTTTTCCCTGTTACTCTATTTGATTGTTGTTCTTCA
TGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCTTTTGATTGTTATT
TTCATAGCATGTTAGTCACTTGAATCAAGCTTTTTCATTTTCAACCAGGG
CAAAAGGTCAAAAGTAACCTACTTTATGAGATCAAAAACAGCAACCCATC
30 GGATAACTTTTAGTTGGAGTTAATAGTTACAATTACCATTGTGATTAATA
ATTATAATATCCTGTATTAATTCATAAAAATTGGTACAGCACATATATGA
CATTTCAAAGGTTTTTGTGTTGACATATATATGCCTCTGGCGTTTTCTTTA
TTGGACTTGCAGACCTCATTCCAAAGTTTATACGGTGACACCTTGGGCCC
TGCTACTTCAGAAGGGACAACCTTGGTCTTTTCATAACTTGATTGAATTAG
35 ATGTGAAATTTAATAAGGATGTTAAAAAGATTATTCCATCCAGTGAGTTG
CTGCAACTGCAAAAGCTGGAAAAGATAAATATAAACAGTTGTGTTGGGGT
AGAGGAGGTATTTGAAACTGCATTGGAAGCAGCAGGGAGAAATGGAAATA
GTGGAATTGGTTTTGATGAATCGTCACAAACAACCTACCACTACTCTTGTG
AATCTTCCAAACCTTAGAGAAATGAACCTTATGGGGTCTAGATTGTCTGAG
40 GTATATATGGAAGAGCAATCAGTGGACAGCATTGAGTTTCCAAAATAA
CAAGAGTTGAAATTAGTAATTGCAACAGTTTAGAACATGTATTTACTAGT
TCCATGGTTGGTAGTCTATCGCAACTCCAAGAGCTACATATAAGTCAGTG
CAAACTTATGGAGGAGGTGATTGTTAAGGATGCAGATGTTTCTGTAGAAG

AAGACAAAGAGAAAGAATCTGATGGCAAGATGAATAAGGAGATACTTGCG
TTACCTAGTCTAAAGTCCCTGAAATTAGAAAGCTTACCATCTCTTGAGGG
GTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATTATTGGATACTTTAA
GAATTGAGGAATGCCCAGCAATAACCACCTTCACCAAGGGAAATTCCGCT
5 ACTCCACAATAAGAGAAATAGAAACAAGATTTGGCTCGGTTTATGCAGG
GGAAGACATCAAATCCTCTATTATAAAGATCAAACAACAGGTAAATCAGA
TCATTGTTGGTTTAATAATTCTTAAACTACATTTGAAAAGTTTCATGTAA
GTTTTTTATTATTGTCAAAAGCCGCAACCTATATTTTCAACTTTATATTT
ATGTACTTTATGCAGGATTTCAAAAAAGCCCAGGACTCTATTTAATGTGA
10 AGTAAATACTAGAAGAGGTAAATTCTATTTACATGTCTCCTGATTGCCTA
TTAATTAATGGCCTTTTCAAGTTCATGGTTTTTGGATGTATTCTTCATGATG
ACGTGAATGTTTAAATACCCCACTAGTTAATTGTTAGGTTGAATGTTGAT
GACCAAAGGACTATATGTCGGGAAGAATATTCAAGGAAAGAATTGTTTCAT
CATATGAAGGGCATTAAATTAAGAAGAACATGGATGCTATGAAGATGTTG
15 GGA.AAATATATGAATCAAATAACAAGCTACTCACTTATCTAAGTTTGTG
GTTGAGGATGTTGATTTTAATATTTCAAATTCATTGGTATCATTATATGG
GTTTATCAGTAGTGTTAATGGGATAATGAGCAACTTAACCTTAAATTATG
CTGTTGGTAAATGTTGGACTCAAGTATGGAAAATTAGGAATAACTTGTGA
AAAATATATGCAAAGTAGGATTGAGATTTTCAATGAAAAAAATTATGAA
20 ACTATACTACTATAGTATATAAATAAATTCAACTTACTGTTGGGTATATT
GGAAGCACATATCATGAAAGTAACTAGAAGCAGAATTTGTTCCCATCTTC
ATCTACTTATAGTTTCCATTTCTTACTTGTA AAAAATCTGATTAAACTTTA
GAGTTATTTCTATTTTTTACCAACCAAATTTTCATATAAAGGCCACAAG
T (SEQ ID NO:107)

25 **RG2J deduced polypeptide sequence (SEQ ID NO:108)**

MSDPTGIVGAIINPIAQTALVPLTDHVGVMISCRKYVRDMQMKMTELNTSRISAEEH
ISRNTRNHLQIPSQIKDWLDQVEGIRANVANFPIDVISCCSLRIRHKLKGQKAFKITEQI
ESLTRQNSLIWTDEPVPLGRVGSMAI STSAASSDHHDVFPSREQIFRKALEALEPVQ
30 KSHIALWGMGGVGKTTMMKKLKEVVEQKKTCNIVQVVIGEKTNP IAIQQA VADY
LSIELKENTKEARADKLRKRFEADGGKNKFLVILDDVWQFFDLEDIGLSPLPNKGV
NFKVLLTSRDSHVCTLMGAEANSILNIKVLKDVEGKSLFRQFAKNAGDDDDLDPAFI
GIADSIASRCQGLPIA KTI ALSLKGRSKSAWDVALSRLENHKIGSEEVVREVFKISYD
NLQDEVTKSIFLLCALFPEDFDIPIEELVRYGWGLKLFIEAKTIREARNRLNNCTERL
35 RETNLLFGSHDFGCVKMHDVVRDFVLHMFSEVKHASIVNHGNMSEWPEKNDTSN
SCKRISLTCKGMSKFPKDINYPNLLILKLMHGDKSLCFPENFYGKMEKVQVISYDKL
MYPLLPSSELCSTNVRVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIG
NLKKLRLLDLTNCKGLRIDNGVLKNLVKLEELYMGVNRPYGQAVSLTDENCNEM
VEGSKLLALEYELFKYNAQVKNISFENLKRFKISVGC SLHGSFSKSRHSYENTLKL
40 AIDKGELLESRMNGLFEKTEVLCLSVGDMYHLSDVVKVSSSFYNLRVLVVSECAEL
KHLFTLGVANTLSKLEHLKVYKCDNMEELIHTGGSEGDTITFPKLLLYLHGLPNL
LGLCLNVNAIELPKLVQMKLYSIPGFTSIYPRNKLEASSLLKEEVVPEELIVEKCGSI
EELFNIDLDCA SVIGEEDNNSSLRNINVENSMKLREVWRIKGADNSRPLFRGFQVVE

KIITRCKRFTNVFTPITTNFDLGALLEISVDCRGNDESDQSNQEQEQIEILSEKETLQE
ATDSISNVVFPSCLMHSFHNLQKLILNRVKGVVFEIESESPTSRELVTTHHNQQQP
VIFPNLQHLDLRGMDNMIRVWKCSNWNKFFTLPKQQSESPFHNLTINIDFCRSIKY
LFSPLMAELLSNLKKVNIKWCYGIEEVVSNRDEDEEMTTFTSTHTTTILFPHLDSL
5 TLSFLENLKCIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAR
EISIEFCNALSSVIPCYAAGQMQLQVLTVSSCNGLKEVFETQLRRSSNKNNEKSGC
DENGGGIPRVNNNVIMLSGLKILEISFCGGLEHIFTFSALESRLQLEELTIMNCWSMK
VIVKKEEDEYGEQQTSTTTTKGTSSSSSSSSSSSSSSSSSSPPSSSKVVPCLKSIVLVNLP
ELVGFFLGMNEFRLPSLDELIEKCPKMMVFTAGGSTAPQLKYIHTRLGKHTIDQES
10 GLNFHQDIYMPLAFSLDLQTSFQSLYGDTLGPATSEGTTWSFHNLIELDVKFNKD
VKKIIPSELLQLQKLEKININSCVGVVEVFETALEAAGRNGNSGIGFDESSQTTTTTL
VNLPNLREMNLWGLDCLRYIWKSNQWTAFEFPKLTRVEISNCNSLEHVFTSSMVGS
LSQLQELHISQCKLMEEVIVKDAADVSEEDKEKESDGKMNKEILALPSLKSLESL
PSLEGFSLGKEDFSFPLDTLRIECPAITTFTKGNSATPQLREIETRFGSVYAGEDIKS
15 SIIKIKQQDFKKAQDSI.CEVNTR

RG2K polynucleotide sequence (SEQ ID NO:109) and (SEQ ID NO:110)

TGGGATTCCATATATAAAAACATATATTTTTATAAAGTGGGATTCCATTG
TTTATATAGATTTTTATTACCAATAGACAATAGATTAAAAAAGATATA
20 AAAACATGTCGGCTTTTGACTAAAAATATAGATTTTTATGAATAGAATAT
TCAATTTGCTTAACTCGTTTAAAAAAATGAAAAAGATGTCGATATAAAA
TCTCATATGGGCCTTCTTTACCATTCAAATAGTAAAATAGTAAAAGATAC
TTGTTTGGGGCATGAACTGACCATAGTCAAACCCATACAAAATCAAACGA
ATCCCACATGGATGATGACGATGGGGTCGCAGTAAATGTGTTTTGGTCCT
25 TTTTTTTCGAGAGAACAGAAGCTTCTGCTCTTCATCTTCTTTAGATTTTG
GGGATTTTCTGGTTTCAGGGGTTTGTGAGTGGAACTAAATTGAAGCAAA
AAAGTATGGTATAATTGGTTGCTAGTGAAATTGATGCTTTCTATTACTAT
CATCTTTAAAATTGTCAAACATTATGTATTAAATTATGAGATCGAAAGT
GGTCTATGGGCCAAAGGTAATACAAGCTTACTCAATGAAATGAATCTAGG
30 ATGCATCATGCATGTATTGGTTAGATTAAAGATTTTCATCAAATTTCTCTT
TATCAAATTGTTGTATACCATGTTATGTAGGTGCTACCACAAGCCATAAC
ATCGAGCAATGGAGTGTATTACTGGCATCTTTAGCAACCCGTTTGCTCAG
TGTCTCATCGCTCCTGTGAAAGAACACCTTTGCCTTCTGATTTTCTATAC
ACAATATGTAGGGGATATGCTTACTGCAATGACGGAGTTGAATGCTGCAA
35 AAGACATTGTTGAAGAGCGGAAGAATCAAAACGTAGAAAAATGTTTTGAG
GTTCCAAACCATGTCAACCGTTGGTTGGAAGATGTTCAAACAATCAACAG
AAAAGTGAACGTGTTCTTAACGATAATTGCAATTGGTTCAATCTATGTA
ATAGGTACATGCTCGCAGTGAAAGCCTTGGAGATAACTCAGGAGATCGAT
CATGCCATGAAACAACCTCTCTCGGATAGAATGGACTGATGATTCAAGTTCC
40 TTTGGGAAGAAATGATTCCACAAAGGCATCCACCTCTACACCATCAAGTG
ATTACAATGACTTCGAGTCAAGAGAACACACTTTTAGGAAAGCACTTGAA
GCACTTGGATCCAACCACACATCCCACATGGTAGCCTTATGGGGGATGGG

TGGAGTTGGGAAGACCACGATGATGAAGAGGCTGAAAAATATTATTAAG
AAAAGAGGACGTTTCATTATATTGTTTTGGTGGTTATAAAGGAAAATATG
GATCTCATTTCCATCCAGGATGCTGTAGCAGATTATCTGGATATGAAGCT
AACAGAAAGCAATGAATCAGAAAGAGCCGATAAACTTCGTGAAGGGTTTC
5 AGGCCAAATCAGATGGAGGTAAGAATAGGTTCCTCATAATACTGGATGAT
GTATGGCAATCTGTTAATATGGAAGATATTGGTTTAAGTCCTTTTCCGAA
TCAAGGTGTCGACTTCAAGGTCTTGTTGACCTCGGAAAACAAAGATGTTT
GTGCAAAAATGGGAGTTGAAGCTAATTTAATTTTCGACGTGAAATTCTTA
ACAGAAGAAGAAGCACAAAGTTTGTTTTATCAATTTGTAAAAGTTTCTGA
10 TACCCACCTTGATAAGATTGGAAAAGCTATTGTAAGAACTGTGGTGGTC
TACCCATTGCCATCAAACCATAGCCAATACTCTTAAAAATAGAAACAAG
GATGTATGGAAGGATGCACCTTCTCGTATAGAGCATCATGACATTGAGAC
AATTGCACATGTTGTTTTTCAAATGAGCTACGACAATCTCCAAAACGAAG
AAGCTCAATCCATTTTTTTTGCTTTGTGGATTGTTTCCTGAAGACTTTGAT
15 ATTCCTACTGAGGAATTGGTGAGGTATGGATGGGGATTGAGAGTATTTAA
TGGAGTGTATACTATAGGAGAAGCAAGACACAGGTGAACGCCTACATCG
AGCTGCTCAAGGATTCTAATTTATTGATTGAAAGTGATGATGTTCACTGC
ATC.AAGATGCATGATTTAGTTCGTGCTTTTGTTTTGGATACGTTTAATAG
ATTCAAGCATTCTTTGATTGTTAACCATGGTAATGGTGGTATGTTAGGGT
20 GGCCTGAAAATGATATGAGTGCCTCATCTTGCAAAAGAATTTCATTAATA
TGC.AAGGGCATGTCCGATTTTCCTAGAGACGTAAAGTTTCCAAATCTCTT
GATTTTGAACTTATGCATGCAGATAAGTCTTTGAAGTTTCCTCAAGACT
TTTATGGAGAAATGAAGAAGCTTCAGGTTATATCATACGATCACATGAAG
TATCCCTTGCTTCCAACATCACCTCAATGCTCCACCAACCTTCGTGTGCT
25 TCATCTTCATCAATGCTCATTGATGTTTGATTGCTCTTCTATTGGAAATC
TGTTGAATCTGGAAGTGCTCAGCTTTGCTAATTCTGGTATTGAGTGGTTG
CCTTCCACAATCGGAAATTTGAAGGAGCTAAGGGTACTAGATTTGACAAA
TTGTGATGGTCTTCGTATAGATAATGGTGTCCCTAAAGAAATTGGTGAAAC
TTG.AAGAGCTTTATATGAGAGTTGGTGGTCGATATCAAAGGCCATTAGC
30 TTC.ACTGATGAAAACCTGCAATGAAATGGCAGAGCGTTCAAAAAATCTTTC
TGC.ATTAGAATTTGAGTTCCTCAAAAACAATGCTCAACCAAAGAATATGT
CATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGATGTTATTTTAAG
GGAGATTTTCGGTAAGATCTTTCACCTCTTTTGAAAACACGTTGCGGTTGGT
CACCAACAGAACTGAAGTCTTGAATCTAGGCTTAATGAGTTGTTTGAGA
35 AAACAGATGTTCTTTATTTAAGTGTGGGAGATATGAATGATCTTGAAGAT
GTTGAGGTAAAGTTGGCACATCTTCCTAAATCCTCTTCCTTCCACAATTT
AAGAGTCCTTATCATTTCTGAGTGTATAGAGTTGAGATACCTTTTCACAC
TTGATGTTGCAAACACTTTGTCAAAGCTTGAGCATCTTCAAGTTTACGAA
TGCGATAATATGGAAGAAATCATACATACAGAGGGTAGAGGAGAAGTGAC
40 AATTACATTCCCAAAGCTGAAGTTTTTATCATTGTGTGGGCTACCAAATC
TGTTGGGTTTGTGTGGTAATGTGCACATAATTAATCTACCACAACCTCACA
GAGTTGAAACTTAATGGCATTCCAGGTTTCACAAGCATATATCCTGAAAA
AGATGTTGAAACATCTAGTTTGTTGAATAAAGAGGTAAATGTGTTTTATG

TTAATACAATACAATCTTTTCAATTAACCGTTTCAAAATATATTGTATGA
TTTATTTTTGTTTGGATGGGGTTATTAATGGGTGATTATTTCTCAGGTTG
TAATTCCTAATTTGGAGAACTTGATATTAGTTATATGAAGGATTTGAAA
GAGATATGGCCTTGTGAATTAGGGATGAGTCAGGAAGTTGATGTTTCTAC
5 GTTGAGAGTGATTAAAGTAAGCAGTTGTGATAATCTTGTGAATCTATTCC
CGTGCAATCCTATGCCATTGATACATCACCTTGAAGAGCTTCAAGTGATA
TTTTGTGGTTCCATTGAAGTGTTATTCAACATTGAGTTGGATTCTATTGG
TCAAATTGGAGAAGGCATCAACAATAGCAGCTTGAGAATCATCCAATTGC
AGA ACTTAGGGAAGCTAAGTGAGGTGTGGAGGATAAAAGGTGCGGATAAC
10 TCTAGTCTTCTCATCAGTGGCTTTCAAGGTGTTGAAAGCATTATCGTTAA
CAAATGCAAGATGTTTAGAAATGTATTCACACCTACCACCACCAATTTTG
ATCTGGGGGCACTTATGGAGATTCGGATACAAGATTGTGGAGAAAAGAGG
AGAAACAACGAATTGGTAGAGAGTAGCCAAGAGCAAGAGCAGGTATGGCT
TTCAATTTCACTTTCTTACTTAATGAAGGATTAAGCTCCTGCTTTTTTGAA
15 TAAAAAGTGGATGAATGACTAAATTCGGGAATGCCACCCGGAAAGTTATC
AACCATTTAGCTACACCATTTTTTGA ACTAATGTTGCAATAAATGCATAA
TATAATTAAAAAATGGTCATTGATAAATGTAAACCAACCTTTTTTATTTA
TTAAAATGTCTACAATAAATGATTTTCTTTATTATATATCATTTTATAAC
AATAAGCTTAAAGATGTTTAAATAGCCAATGTCAGTTATAGATCGTAACT
20 AATTTTTTTATTA ACTAGTTTTAGTTAAGATATCACTCATTATTATTTTTA
TAGAAAAAAGACAAGATTGGCTAATCCTCATAAGAATTTGGAAGATTTAA
GCAAAATATAGAGCTTTTCCAAACATAGCCAATAGTTTCTTTTGCAGGTC
CCATCTACGAAATTATCAATAGATTTGCGATTTTTTTTTTGGCACCCGGGA
AATTTCCATTAAATTA AAAAAAAGTTCAAGCCATTTTGTAGTTGGCACCTG
25 CAAAATGGTAGTTTGCACCTGCGGAAATCACCTTTCACCATTTTCGCATCT
ATGACTTGTGAAAATGTTAATTTGTGAAATGGTCATGTGCACCTCATGAG
AAATACGAAATGGTCAGTAATATGACTTTTTTATATAAATATGATGGTGG
CATATATTTATAGGAAAATATAGCTGCACGATATTAATTAATAGTGAAAT
TAGTTAACTGTATACGATAAGTATACAAAATTTATATGTATGAAGTATAC
30 TCAATTTAGGACGACTCGGGCAATGAAATCATCATTTAATAGGAGCAATG
AAATCATTTTCGAAAAATGTTTACAAATGAATAAAATATTAAATTAACT
TAAAACATTTTGTTAGTAGTTTGAAATTTACAACTGAAATTTGTTGTAT
TTATTAACATTTATAAATGTTGTACTATGATTTTTTTCCTTGTTTGCAAT
ATTCCTTAAAAATCCACCTAAAATCAAATAATTAATCTTTTTCAAGTTG
35 AAAAATGAAAATCGTATGATATAACCGTGTATGGATGTGGAATTATATAT
CAGTTACTAATTACATTTTTTGTGGGATATATGTGCGCAGATTGATATT
GCAATCCCATTCACTCTCACACACTCTTCCAAAACCTCCGTAAACTTGC
TTTGGA AAAAGTATGAAGGAGTGGAGGTGGTGTTTGAGATAGAGAGTCCAA
CAAGTAGAGAATTGATAACAATTCACCATAATCAACAACCACTACTTCCC
40 AACCTTGAGTTATTGGATATAAGTTTTATGGACAGCATGAGTCATGTATG
GAAGTGCAACTGGAATAAATCTTCATTCTTCAAAAACAACAGTCAGAAT
CCCCATTCTGTAATCTCACAACCATA CATATTCAATATTGCCAAAGCATT
AAGTACTTGTTTTCAACTCTCATGGCAAACTTCTTTCCAACCTAAAGAA

GGTCGAGGTAAGAGAGTGTTCATGGTATTGAAGAAGTTGTTTCGAACAGAG
ATGATGAAGATGAGGAAAAGACTACATTTACATCTACATCTTCTGAAAAA
AGCACTAATTTGTTCCCTCGTCTTGAATCTCTCGCTCTTTATCAACTTCC
AAATCTCAAGTGTATTGGTGGTGGTGGTCTGCCAACAGTGGGAACAATG
5 AAATATCTCTTGATAATTCCACTACTACTTCTTTTGTGATCAATCT
AAGGTATGTTTTTTTTTTNGTTNCCCTT (SEQ ID NO:109)

Sequence gap

CCTCCCTAATAATACATGTTATGCACACTATACTAACATATTAGACACGT
AAAGGATAAATGCTATGCCTCATATAATACGTTATATTTATAATCTTTAA
10 ACAATCAAATTTATTAAACAAATAACTAAGTGTGAGCAAAGGCAGGTACC
CGACTAAATTGCCCAAACCAGTCTGGTGGTTCGTGGAATGTTGGGCCAG
GTCGTTAAAACGTCTACACACCGGTTCTTTAAATCACAGATCCGCTTCTC
ATACTGTGAACCCGGTTTTAATTTTAAAAGAAAATTCATTATAAAGTAA
ATGACTTAAACCATTACAAACAACAAAATTTACCATTACAATGTTGGAC
15 TATCATTATTTGCAACATAAACTGAAAATACACATATTTCTTCTGATA
TCAGCATGAGTGGCTGGTGGCTAACCCAAAAATCCATGCATTGTAGATG
TGTGTTACAACACATAGTATCAATGAAAGGCATATTTTAGGCTAGAATT
TAACAATCTGTAATAATATCCCTAAAATAATATCATCATCAACCAACT
AATATAAAACCATTGGGTTTCGTCATTTTAGGTACAAAACATAGATTTTTC
20 TAAGCTTGTTGTATTTAAACATATGCTTTCTAACTTAATTGATTTTGCA
TTCCAAAATTTTAGGTTGTAAAGTGGTATGTCATTTGTTGTCTTTTCAAC
ATTAAATTGTACAAAACCAAACTACATAATTGATGTAGATATCATAACA
ATTGTGTTATTTAGTATATAAAACTAAATTTTGAATTGAATTTCTTATA
CAAAGTTGTGTCTATGTATACATGTTTATGTAGGTAATAGACAATTAGT
25 CTCTGTAAAGTATATGGAGTTTAATTTTLAGACTAATTTTTCATGTGTTG
CAGTTTTATCAGGCAGGTGGCGTTTTTTGGACGTTATGCCAATACTCCAG
AGAGATAAATATAAGGGAGTGTTCATGTCATTGTCAAGTGTAATTCCATGTT
ATGCAGCAGGACAGATGCAAAATGTTCAAGTGCTGAATATATACAGGTGC
AACTCAATGAAGGAGTTATTTGAACTCAAGGGATGAACAACAACATGG
30 TGACAGTGGTTGTGATGAAGGAAATGGTTGTATACCAGCAATTCCAAGAC
TAAATAACGTTATTATGCTACCCAATCTAAAGATATTGAAGATTGAAGAT
TGTGGTCATCTGGAACATGTATTCACATTCTCTGCACTTGGAAGCCTGAG
ACAGCTCGAAGAGTTAACGATAGAGAAATGCAAGGCAATGAAAGTGATAG
TGAAGGAAGAAGATGAATATGGAGAGCAAACAACAAGGCATCTTCGAAG
35 GAGGTTGTGGTCTTTCTCTCGTCTCAAGTCCATTGAACTGGAAAATCTACA
AGAGCTCATGGGTTTCTACTTAGGGAAGAATGAGATTCAGTGGCCTTCAT
TGGATAAGGTTATGATCAAGAATTGCCCAGAAATGATGGTGTTCACCT
GGTGAGTCCACAGTTCCCAAGCGCAAGTATATAAATACAAGCTTTGGCAT
ATATGGGATGGAGGAGGTACTTGAACTCAAGGGATGAACAACAATAATG
40 ATGACAATTGTTGTGATGATGGAAATGGTGGAAATCCAAGACTAAATAAC
GTTATTATGTTTCCAAATATAAAGATATTGCAAATCAGCAATTGTGGCAG
TTTGGAACATATATTCACATTCTCTGCACTTGAAAGCCTGATGCAGCTCA
AAGAGTTAACAATAGCGGATTGCAAGGCAATGAAAGTGATTGTGAAGGAG

GAATATGATGTAGAGCAAACAAGGGTATTGAAGGCTGTGGTATTTTCTTG
TCTAAAGTCCATTACACTATGCCATCTACCAGAGTTGGTGGGTTTCTTCT
TGGGGAAGAATGAGTTCTGGTGGCCTTCATTGGATAAGGTTACCATCATT
GATTGCCCCACAAATGATGGGGTTCACACCTGGTGGGTCAACAACCTCCCA
5 CCTCAAGTACATACTCAAGCTTAGGCAAACATACTCTTGAATGTGGCC
TTAATTTCAAGTCACAACTACTGCATATCATCAGGTATAATTATTATTCT
TTNACACCATCTAATTATGGAATCATGACGCTAATTACAGTATTAAACAC
(SEQ ID NO:110)

10 **RG2K deduced polypeptide sequence (SEQ ID NO:111)**

MECITGIFSNPFAQCLIAPVKEHLCLLIFYTQYVGDMLTAMTELNAAKDIVEERK
NQNVKCFEVPNHVNRWLEDVQTINRKVERVLNDNCNWFNLCNRYMLAVKAL
EITQEIDHAMKQLSRIEWTDDSVPLGRNDSTKASTSTPSSDYNDFESREHTFRKAL
EALGSNHTSHMVALWGMGGVGKTTMMKRLKNIIEKRTFHYIVLVVIKENMDL
15 ISIQDAVADYLDMKLTESNESERADKLREGFQAKSDGGKNRFLIILDDVWQSVN
MEDIGLSFPNPQGVDFKVLLTSENKDVCAKMGVEANLIFDVKFLTEEEAQSLFY
QFVKVSDTHLDKIGKAIVRNCGGLPIAKTIAN TLKNRNKDVWKDALSRIEHHD
IETIAHVVFQMSYDNLQNEEAQSIFLLCGLFPEDFDIPTTEELVRYGWGLRVFNGV
YTIGEARHRLNAYIELLKDSNLLIESDDVHCIMHDLVRAFVLDTFNRFKHS LIV
20 NHGNGGMLGWPENDMSASSCKRISLICKGMSDFPRDVKFPNLLILKLMHADKS
LKFPQDFY GEMKKLQVISYDHMKYPLLPTSPQCSTNLRVLHLHQCSLMFDCSSI
GNLLNLEVL SFANS GIEWLPSTIGNL KELRVLDLTNCDGLRIDNGVLKLVKLEELY
MRVGGRYQKAISFTDENCNEMAERSKNLSALEFEFFKNNAQPKNMSFENLERFKIS
VGCYFKGDFGKIFHSFENTLRLVTNRTEVLESRLNELFEKTDVLYLSVGDMNDLED
25 VEVKLAHL PKSSSFHNLRLVLI SECIELRYLFTLDVANTLSKLEHLQVYECDNMEEII
HTEGRGEVTITFPKLKFLSLCGLPNLLGLCGNVHIINLPQLTELKLNIPGFTSIYPEK
DVETSSLLNKEVVIPNLEKLDISYMKDLKEIWPCELGMSQEVDVSTLRVIKVSSCDN
LVNLFPCNPMPLIHHLEELQVIFCGSIEVLFNIELDSIGQIGEGINSSLRIQLQNLGK
LSEVWRIKGADNSSLLISGFQGVESIIVNKCKMFRNVFTPTTTNFDLGALMEIRIQDC
30 GEKRRNNELVESSQEQEQ

RG2L polynucleotide sequence (SEQ ID NO:112)

GGAAGACACAATGATGCAAAGACTGAAGAAGGTTGCCAAAGAAAATAGAA
TGTTCA GTTACATGGTTCGAGGCAGTTATAGGGGAAAAGACAGACCCAATT
35 GCTATTCAACAAGCTGTAGCCGATTACCTTCGTATACAGTTCAAAGAAAG
CACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAGGCCCACT
CTGNAGACGGTAAGAATAAGTTCCTCGTAATATTTGATGACGTCTGGCAG
TCCGTTGATCTGGAAGATATTGGNTTAAGTCCTTTTCCAAATCAAGGTGT
CGACTTCAAGGTCTTGTTGACTTCACGAGACGAACACGTTTGCACAATGA
40 TGGGGGTTGAAGCTAATTCAGTTATTAATGTGGGACTTCTAACTGAAGTA
GAAGCACAAAGTCTGTTCCAGCAATTTGTAGAACTTTTGAGCCCGAGCT
CTGTAAGATAGGAGAAGTTATCGTAAGAAAGTGTGCGGTCTACCTATTG

CCATCAAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGG
AAGGATGCACTTTCACGTATAGAGCACTATGACATTCGTAGTGTTGCGCC
TAAAGTCTTTGAAACAAGCTATCACAATCTCCAAGACAGGGAGACTAAAT
CCGTGTTTTTGTATGTGTGGTTTGTTCCTGAAGACTTCAATATTCCTACC
5 GAGGAGTTGATGAGGTATGGATGGGGCTTAAAGCTATTTGACAGAGTTTA
TACAATTAGAGAAGCAAGAACCAGGCTCAACACCTGCATTGAGCGACTTG
TGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATG
CATGATCTGGTGCGTGCTTTTGTTTTGGGTATGTATTCTGAAGTCGAGCA
TGCTTCAATTGTCAACCATGGTAATATGCATGGGTGGACTAAAAATGATA
10 TGAACGACTCTTGCAAAACAGTTTCTTTAACATGCGAGAGTGTGTCTGAG
TTTCCAGGAGACCTCAAGTTTCCAAACCTAAAGCTTTTGAAACTTATGCA
TGGAGATAAGATGCTAAGGTTTTCTCAAGACTTTTATGAAGGAATGGAAA
AGCTCCAGGTAATATCATACCATAAAATGAAGTATCCATTGCTTCCCTCG
TCACCTCAATGCTCCACCAACCTTCGAGTGCTTCATCTTCATCGGTGTTC
15 ATTACGGATGCTTGATTGCTCTTGTATCGGAAATTTGACGAATCTGGAAG
TGTTGAGCTTCGCTAATTCTGGCATTGAACGGATACCTTCAGCAATCGGA
AATTTGAAGAAGCTTAGGCAACTTGATCTGAGAGGTCGTTATGGTCTTTG
TATAGAACAGGGTGTCTTGAAAAATTTGGTCGAACTTGAAGAAGTTTATA
TTGGAAATGCATCTGCGTTTAGAGATTATAACTGCAATGAGATGGCAG
20

RG2L deduced polypeptide sequence (SEQ ID NO:113)

EDTMMQRLKKVAKENRMFSYMVEAVIGEKTDPIDAIQQAADYLRIFKESTKPAR
ADKLRWFKAHS?DGKNKFLVIFDDVWQSVLDLEDIGLSPFPNQGVDFKVLTSRDE
HVCTMMGVEANSVINVGLLTEVEAQSLFQQFVETFEPELCKIGEIVVRKCCGLPIAI
25 KTMACTLRNKRKDAWKDALSRIEHYDIRSVAPKVFETSYHNLQDRETKSVFLMCG
LFPEDFNIPTEELMRYGWGLKLFDRVYTIREARLNTCIERLVQTNLLIESDDVGC
VKMHDLVRAFLGMYSEVEHASIVNHGNMHGWTKNDMNDSCKTVSLTCESVSEF
PGDLKFPNLKLLKLMHGDKMLRFSQDFYEGMEKLQVISYHKMKYPLLPSSPQCST
NLRVLHLHRCSLRMLDCSCIGNLTNLEVLFSANGIERIPSAIGNLKKLRQLDLRGR
30 YGLCIEQGVLKNLVELEELYIGNASAFRDYNCNEMA

RG2M polynucleotide sequence (SEQ ID NO:114)

GGGGAAGACACAATAGATGCAAAGGCTGAAGAAGTTGCCAAAGAAAAGAG
AATGTTCAAGTTATATCATTGAGGCGGTTATAGGGGAAAAGACAGACCCCA
35 TTCCATTCAGGAAGCTATATCATATTACCTTGGTGTAGAGCTCAATGCA
AATACTAAGTCAGTAAGAGCTGATATGCTTCGTCAAGGGTTCAAGGCCAA
ATCTGATGTAGGTAAGGATAAATTCTTAATAATACTCGACGATGTATGGC
AGTCTGTTGATTTGGAAGATATTGGATTAAGTCCATTTCCAAATCAAGGT
GTTAACTTCAAGGTCCTGTAAACATCACGAGACCGACATATTTGCACTGT
40 GATGGGGGTTGAAGGTCATTCGATTTTAAATGTGGGACTTCTCACAGAAG
CAGAATCAAAAAGATTGTTCTGGCAGTTTGTAGAAGGTTCTGATCCTGAG
CTCCATAAGATAGGAGAAGATATTGTAAGTAAGTGTTGTGGTCTACCCAT

TGCCATTAAAACCATGGCATGTACACTTAGAGATAAAAGTACGGATGCAT
GGAAGGATGCACTGTCTCGTTTAGAGCATCATGACATTGAAAATGTTGCC
TCTAAAGTTTTTAGAGCGAGCTATGACCATCTCCAAGACGAGGAGACTAA
ATCCACTTTTTTCTATGTGGATTGTTTCCAGAAGATTCCAATATTCCTA
5 TGGAGGAGTTGGTGAGGTATGGGTGGGGATTGAAATTATTTAAAAAAGTG
TATACCATAAGAGAAGCAAGAACTAGGCTCAACACTTGCATTGAGCGGCT
CATCTATACCAATTTGTTGATAAAAGTTGATGATGTTTCAGTGCATCAAGA
TGCATGATCTCATCCGTTCTTTTGTGTTTGGATATGTTTTCTAAAGTTGAG
CATGCTTCGATTGTCAACCATGGTAATACGCTAGAGTGGCCTGCAGATNA
10 TNTGCACGACTCTTGTAAGGGCTTTCATTAAACATGCAAGGGTANATGTG
AGTTTTGTGGAGACCTNAANTTTCCAACCCTAATGATTTTAAACTTATG
CATGGAGATAAATCGCTAAGGTTT

RG2M deduced polypeptide sequence (SEQ ID NO:115)

15 GEDTIDAKAEVAKERMFSYIIEAVIGEKTDPISIQEASYYLGVELNANTKSVRAD
MLRQGFKA KSDV GKDKFLILDDVWQSV DLEDIGLSPFPNQGVNFKVLLTSRDRHI
CTVMGVEGHSIFNVGLL TEAESKRLFWQFVEGSDPELHKIGEDIVSKCCGLPIAKT
MACTLRDKSTDAWKDALSRLEHHDIENVASKVFRASYDHLQDEETKSTFFLCGLFP
EDSNIPMEELVRYGWGLKLFKKVYTIREARTLNTCIERLIYTNLLIKVDDVQCIKM
20 HDLIRSFVLD MFSKVEHASIVNHGNTLEW PAD??HDSCKGLSLTCKG?CEFCGDL?F
PTLMILKLMHGDKSLRF

RG2N polynucleotide sequence (SEQ ID NO:116)

AGGTAAAATCCATAACCCTAAATGTTGGTACGCTCATATATCAAATTGCG
25 TGTTTTGTTGAATGAAAAAAGCATGCTCAAAAAACCAGTGTAAGGCACGG
TATATGACATATTTATAGTTACTGATAACAAATTATGATAATTTTGGGTT
TACRTAAGTTAGGATTTCGTACTTCAACCAAATGTAATAGTTTTTGTGAGT
CTATCTATGTATTTGGGGAATCACATTAGCAACGGGATTGTACTAGTAAT
TCGAAAAAGTCTTTTAAATAATTTTCTGTTTATAATTTATGAATAGTTT
30 TAGCGACATCTAATATTAAATAGAATGTATCTGATATTGAATTAATGTCC
TTAATGTGAACATAGACCTTTTCCATTTACTAATGCCTAATTATTAGTTT
CTAATCAATAAATTTTAATTTCTGTTTTATGCTTCTAAGACAATAAAAAAT
CCATGATTTACCTTTAAATATTAACAAAAATGACCATAAATAAATAAAAA
ATTAGGATACCAAACCCCCCGCCATGCCCAATGTCTAAATATTCTTGAT
35 GCTTTTGCTTTTCCCTCTTTTCCTTGTTAGTCTATTATTCTGGAGAGTTT
GAGAGAGTTTCATACAAGAAAATTTCAAGAAGAAAGCAAAGGTCCAGGTA
TTCTCTTTTCTTAATTATGTATTAACCTTACAAGCATTTTTTACACGATCC
ATGGTTTTTTGTGTATGTTTTTCAAATTGAACTAGATTGGGACTTTTGC
CCTTGATGATTCATAAGATATTGCATGGAGTTGAGATTGTGTAAGAAAAG
40 TGGTGAATAGAAAGAGCAAGTGAATCCAGATATAGTATTGGTAATATATG
ATGATGAGATAGAGATATGTTAAACTGGCTAGAAAATTGTTTTAATTTG
AAATTTAGGTKGTTGAATTTGAAAGATACCAAGCTAATAACTAATTAGTT

ATGCTAAWTAGTTATAAAGAACAACAACTCTTAGTTTTTTTTTTCATGA
TTTTCAACCTCTTTGTACCAAATAAATTATAGCAAAATTGAATATCATT
CTCTGCAATCAATCTTAACTTTTGTTATTATCATCATGTCTAAAATTGCC
ACAAGTTTATTTTCAAAGTCATATTGGATTATGAAAGGACTATTTTACC
5 AATTACATCTTTACTTTATGGGCCAAAGCTAATACAATCCGACTAACTA
AAGGAATATGGGATGCATATAGTTTGCTTCCCGATTATAGATTTCTATCT
AATTTGTCTATTGTACTAATTTAGGTGCCACCACAAGTAAATTTGTTAAA
TGGATATCGTTAATGCCATTCTTAAACCAGTTGTGCGAGACTCTCATGGTA
CCCGTTAAGAAACACATAGGGTACCTCATTTCCCTGCAGGCAATATATGAG
10 GGAAATGGGTATCAAAATGAGGGGATTGAATGCTACTAGACTTGGTGTGCG
AAGAGCATGTGAACCGGAACATAAGCAACCAGCTTGAGGTTCCAGCCCAA
GGCAGGGGTTGGTATGAAGAAGTAGGAAAGATCAATGCAAAAGTGGAAAA
TTTTCTAGCGATGTTGGCAGTTGTTTCAATCTTAAGGTTAGACACGGGG
TCGGAAAGAGAGCCTCCAAGATAATTGAGGACATCGACAGTGTGATGAGA
15 GAACACTCTATCATCATCTGGAATGATCATTCCATTCTTCTAGGAAGAAT
TGATTCCACGAAAGCATCCACCTCAATACCATCAACCGATCATCATGATG
AGTTCCAGTCAAGAGAGCAAACCTTTCACAGAAGCACTAAACGCACTCGAT
CCTAACCACAAATCCCACATGATAGCCTTATGGGGAATGGGCGGAGTGGG
GAAGACGACAATGATGCATCGGCTGAAAAAGGTTGTGAAAGAAAAGAAAA
20 TGTTTAATTTTATTGTTGAGGCGGTTGTAGGGGAAAAAACAGACCCCAT
GCTATTCAATCAGCTGTGGCAGATTACCTAGGTATAGAGCTCAATGAAAA
AACTAAACCAGCAAGAAGCTGAGAAGCTTCGTAAATGGTTTGTGGACAATT
CTGCTGGTAAGAAGATCCTAGTCATACTCGACGATGTATGGCAGTTTGTG
GATCTGAATGATATTGGTTTAAGTCCTTTACCAAATCAAGGTGTGCACTT
25 CAAGGTGTTGTTGACATCACGAGACAAAGATGTTTGCAGTGAAGATGGGAG
CTGAAGTTAATTCAACTTTTAATGTGAAAATGTTAATAGAAACAGAAGCA
CAAAGTTTATTCCACCAATTTGTAGAAATTTCCGGATGATGTTGATCGTGA
GCTCCATAATATAGGAGTGAATATTGTAAGGAAGTGTGGCGGTCTACCCA
TTGTCATCAAAACCATGGCGTGTACTCTTAGAGGAAAAAGCAAGGATGCA
30 TGAAGAATGCACTTCTTCGTTTAGTGAACATAACATTGAAAATATAGT
GAATGGAGTTTTTAAAATGAGTTACGACAATCTCCAAGATGAGGAGACTA
AATCCACCTTTTTGCTTTGTGGAATGTTTCCCGAAGACTTTAATATTCCT
ACCGAGGAGTTGGTGAGGTATGGATGGGGGTTGAAATTATTTAAAAAAGT
GTATACTATAGGAGAAGCAAGAATCAGGCTCAACACATGCATTGAGCGGC
35 TCATTCATACAAATTTGTTGATTGAAGTTGATGATGTTAGGTGCATCAAG
ATGCATGATCTTGTCCGTGCTTTTGTGTTTGGATATGTATTCTAAAGTCGA
GCATGCTTCCATTGTCAACCATGGTAATACTAGAGTGGCATGTGGATA
ATATGCACAACTCTTGTAAGAAAGACTTTCATTAACATGCAAGGGTATGTCT
AAGTTTCTACAGACCTCAAGTTTCCAAACCTCTCGATTTTGAACTTAT
40 GCATGAAGATATATCATTGAGGTTTCCCAAAAACCTTTTATGAAGAAATGG
AGAAGCTTGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCC
TCATCACCGCAATGCTCCGTCAACCTTTGCGTGTTCATCTCCATAAATG
CTCGTTAGTGATGTTTGACTGCTCTTGTATTGGAAATCTGTGCAATCTAG

AAGTGCTTAGCTTTGCTGATTCTGCCATTGACCTGTTGCCTTCCACAATC
GGAATTTTGAAGAAGCTAAGGCTACTGGATTTGACAAATTGTTATGGTCT
TTGTATAGCTAATGGTGTCTTTAAAAAATTGGTCAAACCTTGAAGAGCTCT
ATATGACAGTGGTTAATGGAGGAGTTCGAAAGGCGATCAGCCTCACTGAG
5 GATAACTGCAATGAGATGGCAGAACGTTCAAAAGACCTTTCTGCATTAGA
ACTTGAGTTCTTTGAAAACAATGCTCAGCCAAAGAATATGTCATTTGAGA
AGCTACAACGATTCCAGATCTCAGTGGGGTGCTATTTATATGGAGCTTCC
ATAAAGAGCAGGCACTCGTATGAAAACACATTGAAGTTGGTTATTGACAA
AGGTGAATTATTTGAATCTTGAATGAACGGCCTGTTTAAGAAAACAGAGG
10 TGTTATGTTTAAGTGTGGGAGATATGAATGATCTTGAAGATRTTGAGGTT
AAGTCATCCTCACAACYTCTTCAATCTTCTTCGTTCAACAATTTAAGAGT
CCTTGTCGTTTCAAAGTGTGCAGAGTTGAAACACTTCTTCACACCTGGTG
TTGCAAACACTTTAAAAAAGCTTGAGCATCTTGAAGTTTACAAATGTGAT
AATATGGAAGAACTCATACTAGCAGGGGTAGTGAAGAAGAGACGATTAC
15 ATCCCCAAGCTGAAGTTTTTATCTTTGTGTGGGCTACCAAAGCTATCGG
GTTTGTGCGATAATGTCAAATAATTGAGCTACCACAACCTCATGGAGTTG
GAACTTGACGACATTCCAGGTTTCACAAGCATATATCCCATGAAAAAGTT
TGAAACATTTAGTTTGTGTAAGGAAGAGGTAAATATAAATTTTTAATGCT
AATACATTACAAAGGATCTTTTCAGTTAAATCTTTCAAATATATTGTAA
20 TTTGATTGTATGGGGTATTATTGTTGGATGGGACTATTAATAAATGATTA
TCTTGCAGGTTCTGATTCCTAAGTTAGAGAACTGCATGTTAGTAGTATG
TGG.AATCTGAAGGAGATATGGCCTTGCGAATTTAATATGAGTGAGGAAGT
TAAGTTCAGAGAGATTAAAGTGAGTAACTGTGATAAGCTTGTGAATTTGT
TTCCGCACAAGCCCATATCTCTGCTGCGTCATCTTGAAGAGCTTAAAGTC
25 AAGAATTGTGGTTCCATTGAATCGTTATTCAACATCCATTTGGATTGTGC
TGGTGCAACTGGAGATGAATACAACAACAGTGGTGTAAGAATTATTAAAG
TGATCAGTTGTGATAAGCTTGTGAATCTCTTTCCACACAATCCCATGTCT
ATACTGCATCATCTTGAAGAGCTTGAAGTCGAGAATTGTGGTTCCATTGA
ATCGTTATTCAACATTGACTTGGATTGTGCTGGTGCAATTGGGCAAGAAG
30 ACAACAGAAGCAGCTTAAGAAACATCAAAGTGGAGAATTTAGGGAAGCTA
AGAGAGGTGTGGAGGATAAAAGGTGGAGATAACTCTCGTCCCCTTGTTCA
TGGCTTTCAATCTGTTGAAAGCATAAGGGTTACAAAATGTAAGAGGTTTA
GAAATGTATTACACCTACCACCACAAATTTTAATCTGGGGGCACTTTTG
GAGATTTCAATAGATGACTGCGGAGAAAACAGGGAAAATGACGAATCGGA
35 AGAGAGTAGCCATGAGCAAGAGCAGGTAAGGATTTCAATTTCACTTTCKT
ACTTAATTAATGATTAAGCTCCTGCTTTTTTRAATAAAAAAGGGACAAACC
ATTTTCATGACTTAATGTAGCAATACAAGTCATGTATAAGAGTGACCAACT
CTTTTTTATTTATAAAATGACTACAAAATATTTTTTTTCATTAGAGATCA
TGTATAAATGTGACTAATTTTTTCATCACCTAACTTTAGTTGATAAATCTT
40 TATAAATGTCACTAGTTACTTTTCAGTAAAATAACAAATTTAATAAATTA
TCAACAAAAAGCATCAACTAAAAAATCCCAACACCCGTAATAATTTAAA
ATAAAAGGATTTAACATCTAATACGAACAATTTTTTTTCTAAACATGATT
TGGACCAATATCACCAGCAACTCAAGTTTGGAATCGATTCAGCTTAAAA

CTTGACCARCATAATTAGATAGATGAGAGTTGAAGCTAAAGTGCCTATAT
AAGTTCGTTTCATCTTTTTTCTTGATCTTGATAGCAAGTTGAATSATTTT
CTTCTTCAAAATTGATAAAAATCTACATTATAAAGAGACTAGCTTGAAAA
AAAATGGTCTAGGTGGGTCTTGGGTCTGGTAGATGAAGATGGAAGGGAGA
5 GTAGATTTCAAAGACACAAACACATCTTCATTTTATTTATTTATTTATTA
TTATTATTTTTTGATATCTTGCTCATATTTGTTACAGATATGTGAGGTCT
ATTAATCTTTTTTAAATATATAAAAAATAAATACATAAATGAGAAAATTAA
ATAAAGAATAAATTAATAAGGGCACAATAGTCTTTTTTGGTAAGACAAGG
ACCAAAAGCGCAACAAAAGTAAACAGTAGGGACCATCCGATTTAAAAAAT
10 TAATTAGGGACCAAAAACATAAATTCCCCCAAACCATAGGGACCATTCGT
GTAATTTACTCTTGCTTTTTCGTTTTTGTTTCATATTTGGGTAACATTTTTT
TTGTACATATCTAGGTAACGAACTTGTTGAAAGTGTTACATCTACGATG
TGACCTACTACAACCGATCATAATGGTCATATATGAACACTTCCAACAAG
TTTGTTATCTAGGTGTGTACAAAAAACGATAGTTACCATGATGTGAACA
15 TACCAAAAAATTAATTACCTTAGCAAGTTATTTTCCCATTTAGGTTGTAT
GGAAACAGTTCCGTGAGACCGTGACTTGATGGTAGATAAATTTAGTAA
CTTAACCCTTCAATTAACCTACCTTTTTCTTATTAACCTCAATTTCAAGCT
AAATTCTGATTCTTGTTTGAAAGTAAGTTGCATCTTTATGTTTGTATTAT
CTTGTTGCATAGGATCCTTAGCATCTTTTAATAATTTATTTGAAGGTGAA
20 AGATCCAACCTATTTTTAATCTGTTGGCATTTCATCATTTGCAACTGTT
TCTTGAAAAAAA::TACCTAAAATCAAATAACCATTTTCATATCCAAAA
TTATAAGAGAGAATTGTTAACGGACATGGAATCATAAATCATTAACACAG
TTCAGTACACAGGTTGCTAATTACATTTCTTGCTGTGCAGATTGAAATTC
TATCAGAGAAAGAGACATTACAAGAAGCCACTGGCAGTATTTCAAATATT
25 GTATTCCCATCCTGTCTCATGCACTCTTTTCATAACCTCCATAAACTTAA
CTTGAACAGAGTTGAAGGAGTGAGGTGGTGTGTTGAGATAGAGAGTGAGA
GTCCAACAAGTAGAGAATTGGTAACAACCTCACCATAACCAACAACAACCT
ATTATACTTCCCAACCTCCAGGAATTGATTCTATGGAATATGGACAACAT
GAGTCATGTGTGGAAGTGCGGCAACTGGAATAAATTCTTCACTCTTCCAA
30 AAGAACAATCAGAATCCCCATTCCACAACCTCAGTAACATACATATTTAT
GAATGCAAAAGCATTAAAGTACTTGTTTTACCTCTCATGGCAGAACTTCT
TTCCAACCTAAAGCATATCGAGATAAGAGAGTGTGATGGTATTGAAGAAG
TTGTTTCAAAAAGAGATGGTGAGGATGAAGACATGACTACATCTAC:::
:::GCACACAACCACCTTTTTCCCTCATCTTGATTCTCTCACTCTAAA
35 GCAACTGAAGAATCTGAAGTGTATTGGTGGAGGTGGTGCCAAGGATGAGG
GGAGCAATGAAATATCTTTCAATAATACCACTGCAACTACTGCTGTTCTT
GATCAATTTGAGGTATGCTTTGTACATATTCAATTATTTATTTAATTTCC
TTGTTAATTTCTTTTTTCTTTGCAATATTCTATGAAAAAATCACCAAA
TCACAAATAAGAGATTTAACTTTTATTTACACCCATGCGGACTCAAGA
40 ATGGGATTTGGAGGCATATAAAGTTACATTCATTTGAACAAGTATTACCA
TTT.ATTTGTTATTTATCATTTTCATATCATTTACTGATAACATTTCTTTT
TTACTTTTCTAATTAGAAAAGGTCCACATGTCTAATTAGGTTTTCCATTC
TATGTGAATCCTCTATTCTGTCTGTAATCAAGCATCTTAGATTATTTATC

CATTTTCATAATTGTGTTTATATTGACAGTTTTTTTCTTTTTATAGTTGT
AATTGCAACCTGTCATATWTTMWWKWCWWATKYWMWWARTAATACATTT
TATACCCWCTATACTAAGATA

5 **RG2N deduced polypeptide sequence (SEQ ID NO:117)**

LGKTTMMHRLKKVVKEKKMFNFIVEAVVGEKTDPIAIQSAVADYLGIELNEKTKPA
RTEKLRKWFVDNSAGKKILVILDDVWQFVDLNDIGLSPLNQGVDFKVLLTSRDKD
VCTEMGAEVNSTFNVKMLIETEAQSLFHQFVEISDDVDRELHNIGVNIVRKCGGLPI
VIKTMACTLRGKSKDAWKNALLRLVNYNIENIVNGVFKMSYDNLQDEETKSTFLL
10 CGMFPEDFNIPTEELVRYGWGLKLFKKVYTIGEARIRLNTCIERLIHTNLLIEVDDVR
CIKMHDLVRAFVLDMYSKVEHASIVNHGNTLEWHVDNMHNSCKRLSLTCKGMSK
FPTDLKFPNLSILKLMHEDISLRFKPNFYEEMEKLEVISYDKMKYPLLSPSPQCSVNL
CVFHLHKCSLVMFDCSCIGNLSNLEVL SFADSAIDL LPSTIGILKKLRLDLTNCYGL
CIANGVFKKL VKLEEL YMTVVNGGVRKAISL

15

RG2O polynucleotide sequence (SEQ ID NO:118)

TTGTAAAACGACGGCCAGTCGAATCGTAACCGTTCGTACGAGAATCGCTG
TCCTCTCCTTCATTTGAATCATGATATTTGAATATCGATACTTTTGACTG
TAGCTTTTGGGTCGATTTTTTAGCAAGATACATAACTGGCCAAACCCATT
20 GGCTATTTTAGCCCAAATATGAAATGGACTGGATTGTTTTTTTCCTTTC
TAACACGCACACATCTGGCGATCAGTATCACTCCATTATGAAGACCTAGT
CAAATTCATTAACGTTTCAGTCGTTCCCTTCAAAGTTTCAAAGTTCCAATT
CCAATTCCCTCTTTTTTTTTTTCTTTCCTCGATTCTGATTTGAATCCGAT
TCTGCGACGAAGGAGAGCTTGGTCAGAGGGCTGTGATTCTTGAGTCTTGA
25 CCTCCGAATCTAGCTGGATTATTTTCGACACACCAGACCACGTATCAGGT
TGCTCATCCCGAAATACTGCTTTGCAAAGTGTGTATCATCGCCTAGGAA
ATT.AAGTTTCTTTTTTTGGCTCTGTTACTGAATCAGTAGCTTTGCAACTTG
CTCATTATAAGCTGATCCATATTTTACATATCTTTTGAAGAATAATAGGT
ACTGACTTTACCTTTCTGATGAGAGCGATTAAAGAGATACCTCTGTAAAA
30 TCCATTTTTGTGAAGGGATCTGGGTTAGTTTTTAAAGGATTTGCTACAAC
AGTATCCCAACAAACGATCTATTTCCCATTTNACTCATCCGCTCAAGATCT
ATCCACCTTTATATATGTTAATTGGGAGTCTTCCATGGTGCAATGAATCT
AGGATGCATTTAGAAGCCCAATCCATTACAAGTTTTCATCCAATTTTCATG
TGACAAGTTGTTGGTTACTATGTAGGTACTTCCACAATTAAGAATTTCCA
35 GCAATGGATGTTGTTAATGCCATTCTTAAACCAGTTGCCGAGACACTTAT
GGAACCTGTTAAGAAACATCTAGGCTACATCATTTCCAGCACAAAACATG
TGAGGGATATGAGTAACAAAATGAGGGAGTTGAACGCTGCAAGACATGCT
GAAGAAGACCACTTGGACAGGAACATAAGAAGTTCGTCTTGAGATTTCAA
TCAAGTTAGGAGTTGGTTAGAAGAAGTAGAAAAGATCGATGCAAAAGTAA
40 AAGCCCTTCCTAGTGATGTCACCGCTTGTTGCAGTCTCAAGATCAAACAT
GAAGTCGGAAGGGAAGCCTTGAAGCTAATTGTGGAGATTGAAAGTGCCAC
AAGACAACACTCTTTGATCACCTGGACTGATCATCCCATTCCTCTGGGAA

AAGTTGATTCCATGAAGGCATCGATGTCCACAGCATCAACCGATTACAAT
GACTTTCAGTCAAGAGAAAAAACTTTTACTCAAGCATTGAAAGCACTTGA
ACCAAACAACGCTTCCCACATGATAGCGTTATGTGGGATGGGTGGAGTGG
GGAAGACCACAATGATGCAAAGACTAAAAAAAGTTGCTAAACAAAATAGA
5 ATGTTTCAGTTATATGGTTGAGGCAGTTATAGGGGAAAAGACGGACCCAAT
TGCTATTCAACAAGCTGTAGCGGATTACCTTCGTATAGAGTTAAAAGAAA
GCACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAGGCCAAC
TCTGGAGAAGGTAAGAATAAATTCCTTGTAATACTTGATGACGTCTGGCA
GTCTGTTGATCTAGAAGATATTGGTTTAAGTCCTTTTCCAAATCAAGGTG
10 TCGACTTCAAGGTCTTATTGACTTCACGAGACGAACATGTTTGACAGTA
ATGGGAGTTGGATCTAATTCAATTCTTAATGTGGGACTTCTAATAGAAGC
AGAAGCACAAAGTTTGTTCCAACAATTTGTAGAACTTCTGAGCCCGAGC
TCCATAAGATAGGAGAAGATATTGTAAGGAAGTGTTGCGGTCTACCTATT
GCCATCAAAACCATGGCATGTACTCTTAGAAATAAAAGAAAGGATGCTTG
15 GAAGGATGCACTTTCGCGTATAGAGCACTATGACCTTCGCAATGTTGCGC
CTAAAGTCTTTGAAACGAGCTACCACAATCTCCATGACAAAGAGACTAAA
TCAGTGTTTTTGATGTGTGGTTTGTTTCCGGAAGACTTCAATATTCCTAC
TGAGGAGTTGATGAGGTATGGATGGGGATTAAAGATATTTGATAGAGTCT
ATACATTTATAGAAGCAAGAAACAGGATCAACACCTGCATTGAGCGACTG
20 GTGCAGACAAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGAT
GCATGATCTGGTCCGTGCTTTTGTTTTAGGTATGTATTCTGAAGTAGAGC
ATGCTTCAGTTGTCAACCATGGTAATATACCTGGATGGACTGAAAATGAT
CCG.ACTGACTCTTGTAAGCAATTTCAATTAACATGCGAGAGTATGTCTGG
AAACATTCCAGGAGACTTCAAGTTTCCAAACCTAACGATTTTGAAACTTA
25 TGCATGGAGATAAGTCGCTAAGATTTCCACAAGACTTTTATGAAGGAATG
GAAAAGCTCCAGGTTATATCATACGATAAAATGAAGTATCCAATGCTTCC
CTTGTCTCCTCAATGCTCCACCAACCTTCGAGTGCTTCATCTCCATGAAT
GTTCAATAAAGATGTTTGATTGCTCTTGTTATTGGAAATATGGCGAATGTG
GAAGTGTTGAGCTTTGCTAATTCTGGCATTGAAATGTTACCTTCCACTAT
30 CGGAAATTTAAAGAAGCTAAGGTTACTTGATTTAACAGATTGTCATGGTC
TTCATATAACACACGGTGTCTTTAACAATTTGGTCAAACCTTGAAGAGTTG
TATATGGGATTTTCTGATCGACCTGATCAAACCTCGTGGTAATATTAGCAT
GACAGATGTCAGCTACAATGAATTAGCAGAACGTTCAAAAGGCCTTTCTG
CATTAGAGTTCCAGTTCTTTGAAAACAATGCCCAACCAAATAATATGTGCG
35 TTTGGGAAACTTAAACGATTCAAGATCTCAATGGGATGCACTTTATATGG
AGGATCAGATTACTTTAAGAAAACGTATGCTGTCCAAAACACATTGAAGT
TGGTTACTAACAAAGGTGAACTATTGGACTCTAGAATGAACGAGTTGTTT
GTTGAAACAGAAATGCTTTGTTTAAGTGTTGATGATATGAATGATCTTGG
TGATGTTTGTGTGAAGTCCTCACGTTCTCCTCAACCTTCTGTGTTCAAAA
40 TTCTAAGAGTCTTTGTGCTTTCCAAGTGTTGAGTTGAGATACCTTTTC
ACAATTGGTGTAGCCAAGGATTTGTCAAATCTTGAGCATCTTGAAGTTGA
TTCATGTAATAATATGGAACAACTCATATGTATTGAGAATGCTGGAAAAG
AGACAATTACATTCCTAAAGCTGAAGATTTTATCTTTGAGTGGGCTACCA

AAGCTTTCGGGTTTGTGCCAAAATGTCAACAACTTGAGCTACCACA
CATAGAGTTGAACTTAAGGGCATTCCAGGGTTCACATGCATTTATCCGC
AAAACAAGTTGGAAACATCTAGTTTGTGTAAGGAAGAGGTAGATATATGT
TTTATGTTAATACAAGTTAAAAAATCTTTTAACTAAAAGTTTCAGTATA
5 TATATCTATATGTCTATAATTTGATTATATGATGTATTAGTGTTTGGATG
TGGCTATTAAGGGATGATTATTTTGCAGGTTGTGATTCCTAAGTTGGAGA
CACTTCAAATTGATGAGATGGAGAATTTAAAGGAAATATGGCATTATAAA
GTTAGTAATGGTGAGAGAGTTAAGTTGAGAAAGATTGAAGTGAGTAACTG
TGATAAGCTTGTGAATCTATTTCCACACAACCCCATGTCTCTGCTGCATC
10 ATCTTGAAGAGCTTGAAGTCAAGAAATGTGGTTCCATTGAATCGTTATTC
AACATCGACTTGGATTGTGTTGATGCCATAGGAGAAGAAGACAACATGAG
GAGCTTAAGAAACATTAAAGTGAAGAATTCATGGAAGTTAAGAGAAGTG
GGTGTATAAAAGGTGAAAATAACTCTTGCCCCCTTGTTTCTGGCTTTCAA
GCTGTTGAAAGCATAAGCATTGAAAGTTGTAAGAGGTTTAGAAATGTATT
15 CACACCTACCACCACCAATTTTAATATGGGGGCACTTTTGGAGATATCAA
TAGATGACTGTGGAGAATACATGGAAAATGAAAAATCGGAAAAGAGTAGC
CAAGAGCAAGAGCAGGTATGGATTTCAATTTCACTTTCTTACTTACTTAA
GGATTAAGCTTCTGTTTTTTTGAATAAAAAAGGGACATCTTCTAATAATG
CACATCTTAAATTAAAAAGTATTTAATTGTTGCATAGCAGCGTATAACAT
20 CTTCTAATAATTTATCTGAAGGTGAAAGATCCAATACTTCTAATTTGTT
AACAAATTTCAATCATTTGCAAATGTTCCCTTAAAAAATTAATTACCTGAAA
TCAAAACAATCTTCTTCAAATCCAAAATTATGAGACAGAATTGAGAAGGG
ATGTGAAATTATAAACCATTAAACACAATTCCATGCTCACGTTACTAATTA
CATTTCTTGTTGGGATATATATGTACAGACTGATATTTTGTGAGAGGAAG
25 TGAAATTACAAGAAGTCACTGATACTATTTCTAATGTTGTATTACATCG
TGTCTCATACACTCTTTTTATAACAACCTCCGTAACTCAACTTGGAGAA
GTATGGAGGAGTTGAGGTTGTGTTTGAAGATAGAGAGTTCAACAAGTAGAG
AATTGGTAACAACATACCATAAACAACAACAACAACAACCTATATTT
CCC.AACCTTGAGGAATTATATCTATATTATATGGACAACATGAGTCATGT
30 ATGGAAGTGCAACAACCTGGAATAAATTTTACAACAATCAGAATCCCAT
TCC.ACAACCTCACAACCATACACATGTCCGATTGCAAAGCATTAAAGTAC
TTGTTTTACCTCTCATGGCAGAACTTCTTTCCAACCTAAAGAGAATCAA
TATTGACGAGTGTGATGGTATTGAAGAAATTGTTTCAAAAAGAGATGATG
TGG.ATGAAGAA
35

RG2O deduced polypeptide sequence (SEQ ID NO:119)

MDVVNAILKPVAETLMEPVKKHLGYIISSTKHVRDMSNKMRELNAARHAEEDHLD
RNIRTRLEISNQVRSWLEEEVEKIDAKVKALPSDVTACCSLKIKHEVGREALKLIVEIE
SATRQHSLITWTDHPIPLGKVDMSMKASMSTASTDYNDFQSREKTFTQALKALEPNN
40 ASHMIALCGMGGVGKTTMMQRLKKVAKQNRMFSYMVEAVIGEKTDPIDAIQQAVA
DYLRIELKESTKPARADKLREWFKANS GEGKNKFLVILDDVWQSVLEDIGLSPFP
NQGVDFKVLLTSRDEHVCTVMGVGSNSILNVGLLIEAEAQSLFQQFVETSEPELHKI

GEDIVRKCCGLPIAIKTMACTLRNKRKDAWKDALSRIEHYDLRNVAPKVFETSYHN
LHDKETKSVFLMCGLFPEDFNIPTEELMRYGWGLKIFDRVYTFIEARNRINTCIERL
VQTNLLIESDDVGCVKMHDLVRAFVLGMYSEVEHASVVNHGNIPGWTENDPTDSC
KAISLTCESMSGNIPGDFKFPNLTILKLMHGDKSLRFPQDFYEGMEKLQVISYDKMK
5 YPMLPLSPQCSTNLRVLHLHECSLKMFDSCIGNMANVEVLSFANSIEMIPLSTIGN
LKKLRLLDLTDCHGLHITHGVFNNLVKLEELYMGFSDRDPQTRGNISMTDVSYNE
LAERSKGLSALEFQFFENNAQPNNMSFGKLRFKISMGCTLYGGSDYFKKTYAVQ
NTLKLVTNKGELLSRNMNELFVETEMLCLSVDDMNDLGDVCVKSSRSPQPSVFKIL
RVFVVSKCVELRYLFTIGVAKDLSNLEHLEVDSNNMEQLICIENAGKETITFLKLKI
10 LSLSGLPKLSGLCQNVNKLLELPQLIELKLKGIPGFTCIYPQNKLETSSLLKEEVVIPKL
ETLQIDEMENLKEIWHYKVSNGERVKLRKIEVSNCDKLVNLFPHNPSLLHHEEL
EVKKCGSIESLFNIDLDCVDAIGEEDNMRSRLNIKVKNSWKLREVWCIKGENNSCPL
VSGFQAVESISIESCKRFRNVFTPTTTNFMGALLEISIDDCGEYMENЕКSEKSSQEQ
EQTDILSEEVKLQEVTDITISNVVFTSCLHSFYNNLRKLNLEKYGGVEVVFEIESSTS
15 RELVTTYHKQQQQQQPIFPNLEELYLYMDNMSHVWKCNNWNKFLQQSESPFHN
LTTIHMSDCKSIKYLFSPLMAELLSNLKRINIDECDGI

RG2P polynucleotide sequence (SEQ ID NO:120)

CCCATTGCTATTCAGGAAGCAGTAGCAGATTACCTCNGTATAGAGCTCAA
20 AGAAAAAACTAAATCNGCAAGAGCTGATATGCTTCGTAAAATGTTAGTTG
CCAAGTCCGATGGTGGTAAAAATAAGTTCCTAGTAATACTTGACGATGTA
TGGCAGTTTGTGTGATTTAGAAGATATCGGTTTAAGTCCTTTGCCAAATCA
AGGTGTTAACCTCAAGGTCTTGCTAACATCACGGGATGTAGATGTTTGCA
CTATGATGGGAGTCGAAGCCAATTCAATTCTCAACATGAAAATCTTACTA
25 GATGAAGAAGCACAAAGTTTGTTCATGGAGTTTGTACAAATTTTCGAGTGA
TGTTGATCCCAAGCTTCATAAGATAGGAGAAGATATTGTAAGAAAGTGTT
GTGGTTTGCCTATTGCCATCAAAACCATGGCCCTTACTCTTAGAAATAAA
AGCAAGGATGCATGGAGTGATGCACTTTCTCGTTTAGAGCATCATGACCT
TCACAATTTTGTGAATGAAGTTTTTGAATTAGCTACGACTATCTTCAAG
30 ACCAGGAGACTAAATATATCTTTTTGCTTTGTGGATTGTTTCCCGAAGAC
TACAATATTCCTCCTGAGGAGTTAATGAGGTATGGATGGGGCTTAAATTT
ATTTAAAAAAGTGTATACTATAAGAGAAGCAAGAGCCAGACTCAACACCT
GCATTGAGCGGCTTATCCATACCAATTTGTTGATGGAAGGAGATGTTGTT
GGGTGTGTAAAGATGCATGATCTAGCACTTGCTTTTGTATGGATATGTT
35 TTCTAAAGTGCAAGATGCTTCAATTGTCAACCATGGTAGCATGTCAGGGT
GGCCTGAAAATGATGTGAGTGGCTCTTGCCAAAGAATTCATTAACATGC
AAGGGTATGTCTGGGTTTCCTATAGACCTCAACTTCCAAACCTCACAAT
TTTAAACTTATGCATGGAGATAAGTTTCTCAAGTTTCCTCCAGACTTTT
ATGAACAAATGGAAAAGCTTCAAGTTGTATCGTTTCATGAAATGAAATAT
40 CCGTTTCTTCCCTCGTCTCCTCAATATTGCTCCACCAACCTTCGAGTTCT
TCATCTCCATCAATGCTCATTGATGTTTGATTGCTCTTGTATTGGAAATC
TGTTTAATCTGGAAGTGTTGAGCTTTGCTAATTCTGGCATTGAATGGTTA

5 CCTTCCAGAATTGGAAATTTGAAGAAGCTAAGGCTACTAGATTTGACAGA
TTGTTTTGGTCTTCGTATAGATAAGGGTGTCTTAAAAAATTTGGTCAAAC
TTGAAGAGGTTTATATGAGAGTTGCTGTTCGAAGCAAAAAGCCGGAAAT
AGAAAAGCCATTAGCTTCACAGATGATAACTGCAATGAGATGGCAGAGCG
TTC

RG2P deduced polypeptide sequence (SEQ ID NO:121)

PIAIQEAVADYL?IELKEKTKSARADMLRKMLVAKSDGGKNKFLVILDDVWQFVDL
EDIGLSPLPNQGVNFKVLLTSRDVDVCTMMGVEANSILNMKILLDEEAQSLFMEFV
10 QISSDVPKLHKIGEDIVRKCCGLPIAKTMALTLRNKSKDAWSDALSRLHHDLHN
FVNEVFGISYDYLQDQETKYIFLLCGLFPEDYNIPPEELMRYGWGLNLFKKVYTIRE
ARARLNTCIERLIHTNLLMEGDVVGCVKMHDLLALAFVMDMFSKVQDASIVNHGS
MSGWPENDVSGSCQRISLTCKGMSGFPIDLNFPNLTKLMLHGDKFLKFPPDFYEQ
MEKLQVVVSFHEMKYPFLPSSPQYCSTNLRVLHLHQCSLMFDCSCIGNLFNLEVLSF
15 ANSGIEWLPSRIGNLKKLRLDLLTDCFLRIDKGVLKNLVKLEEVYMRVAVRSKKA
GNRKAISFTDDNCNEMAERS

RG2Q polynucleotide sequence (SEQ ID NO:122)

TGGGGAAGACACAGTGATAGAAAARAAAAGAATGTTGTGGAAAAGAGGA
20 AAATGTTTGATTATGCTGTTGTGGCGGTTATAGGGGAAAAGACGGACCCT
ATTGCTCTTCAGAAAAGTGTGCGGATTACTTGCATATTGAGCTAAATGA
AAGCACTAACTAGCAAGAGCAGATAAACTTTGCAAATGGTTCAAGGACA
ACTCGGATGGAGGTAAGAAAAAGTTCCTCGTAATACTCGACGATGTTTGG
CAATCTGTTGATTTGGAAGATATTGGTTTAAGTACTCCTTTTCCAAATCA
25 AGGTGTCAACTTCAAGGTTTTGTTGACATCACGAAAGAGAGAAATTTGCA
CAATGATGGGAGTTGAAGCTGATTTAATTCTCAATGTCAAAGTCTTAGAA
GAAGAAGAAGCACAAAAGTTGTTCCCTCCAGTTTGTAGAAATTGGTGACCA
ATACCACGAGCTTCATCAGATAGGGGTACATATAGTAAAGAAGTGTTATG
GTTTACCCATTGCCATTAAAACCATGGCTCTTACTTTAAGAAATAAAAGA
30 AAGGATTCATGGAAGGACGCACTCTCTCGTTTAGAGGACCATGACACTGA
AAATGTTGCAAATGCAGTTTTTCGAGATGAACTACCGCAATCTACAAGATG
AGGAGACCAAAGCCATTTTTTTTGCTTTGCGGTTTGTTCCTCCGAAGACTTT
GATATTCCTACTGAGGAGTTGGTGAGGTATGGATGGGGCTTAAATCTATT
TAAAAAAGTGTATACCATAAGAAAGGCAAGAACGAGATCGCATACATGTA
35 TTGAGCGACTCTTGGATTCAAATTTGTTGATTGAAAGTAACGATATTCGG
TGCGTCAAGATACACGATCTGGTGCGCGCTTTTGTGTTTGGATATGTATTG
TAAAGTTGAGCATGCTTCAATTGTCAACCATGGTAATATGCGGACCGAAT
ATAATATGGCTGACTCTTGCAAAACAATTTCAATTAACATACAAGAGTATG
TCTGGGTTTGAAGTTTCCAGGAGACCTCAAGTTTCCAAACCTAACAGTTTT
40 GAACTTATGCANGGAGATAAGTCTCTAAGGTTTCCTCAAGACTTTTATC
AATCAATGGAAAACTTCGGGTATATCATATGATAAAATGAAGTATCCA
TTGCTTCCCTCATCACCTCAATGCTCCACTAACATCCGAGTGCTTCGTCT

CCATGAATGTTTCATTAAGGATGTTTGATTGCTCTTGTATTGGAAAGCTAT
TGAATTTGGAAGTCCTCAGCTTTTTTAATTCTAACATTGAATGGTTACCT
TCCACAATCAGAAATTTAAAAAAGCTAAGGCTACTAGATTTGAGATATTG
TGATCGTCTTCGTATAGAACAAGGTGTCTTGAAAAATTTGGTCAAACCTG
5 AAGAACTTTATACTGGATATACATCAGCGTTTACAGA

RG2Q deduced polypeptide sequence (SEQ ID NO:123)

GEDTVIEKKKNVVEKRKMFDYAVVAVIGEKTDPIALQKTVADYLHIELNESTKLAR
ADKLCKWFKDNSDGGKKKFLVILDDVWQSVLEDIGLSTPPNQGVNFKVLLTSR
10 KREICTMMGVEADLILNVKVLLEEEEAQKLFLQFVEIGDQYHELHQIGVHIVKKCYG
LPIAIKTMALTLRNKRKDSWKDALSRLEDHDTENVANAVFEMNYRNLQDEETKAI
FLLCGLFPEDFDIPTEELVRYGWGLNLFKKVYTIRKARTRSHTCIERLLDSNLLIESN
DIRCVKIHDLVRAFVLDMYCKVEHASIVNHGNMRTEYNMADSCKTISLTYKSMMSG
FEFPGDLKFPNLTVLKLM?GDKSLRFPQDFYQSMEKLRVISYDKMKYPLLPSSPQCS
15 TNIRVLRLHECSLRMFDCSCIGKLLNLEVLSSFNSNIEWLPSTIRNLKKLRLLDLRYC
DRLRIEQGVKLNLVKLEELYTGYTSAFTE

RG2S polynucleotide sequence (SEQ ID NO:124)

ATTTGGGGTTTTACATTTAATTTTTTTGTGCATGAATGTGAAAATAGACTG
20 CTTATTGATTCTTTGTGTTTCATTGAGTTGATTTTCATTATTACTACCTT
ACA.AATTGCTCAGTGATAGATTTCCATTAATTTGCTAATTCGGTTGCTTC
TAA.ATATGTAGGAGCTACTAAAAGCAAAAATATCGAGCAATGTCGGACCC
AACGGGGATTGCTGGTGCCATTATTAACCCAATTGCTCAGAGGGCCTTGG
TTCCCGTTACAGACCATGTAGGCTACATGATTTCTGCAGAAAATATGTG
25 AGGGTCATGCAGACGAAAATGACAGAGTTGAATACCTCAAGAATCAGTGT
AGAGGAACACATTAGCCGGAACACAAGAAATCATCTTCAGATTCCATCTC
AAATTAAGGATTGGTTGGACCAAGTAGAAGGGATCAGAGCAAATGTGGAA
AACTTTCCGATTGATGTCATCACTTGTTGTAGTCTCAGGATCAGGCACAA
GCTTGGACAGAAAGCCTTCAAGATAACTGAGCAGATTGAAAGTCTAACAA
30 GACAGCTCTCCCTGATCAGTTGGACTGATGATCCAGTTCCTCTAGGAAGA
GTTGGTTCCATGAATGCATCCACCTCTGCATCATCAAGTGATGATTCCC
ATCAAGAGAGAAAACTTTTACACAAGCACTAAAAGCACTCGAACCCAACC
AAC.AATTCCACATGGTAGCCTTGTGTGGGATGGGTGGAGTAGGGAAGACT
AGAATGATGCAAAGGCTGAAGAAGGCCGCTGAAGAAAAGAAATTGTTTAA
35 TTATATTGTTAGGGCAGTTATAGGGGAAAAGACGGACCCCTTTGCCATTC
AAGAAGCTATAGCAGATTACCTCGGTATACAACCTCAATGAAAAAACTAAG
CCAGCAAGAGCTGATAAGCTTCGTGAATGGTTCAAAAAGAATTCAGATGG
AGGTAAGACTAAGTTCCTCATAGTACTTGACGATGTTTGGCAATTAGTTG
ATCTTGAAGATATTGGGTAAAGTCCTTTTCCAAATCAAGGTGTCGACTTC
40 AAGGTCTTGTTGACATCACGAGACTCACAAGTTTGCACATGATGGGGGT
TGAAGCTAATTCAATTATTAACGTGGGCCTTCTAACTGAAGCAGAAGCTC
AAAGTCTGTTCCAGCAATTTGTAGAACTTCTGAGCCCGAGCTCCAGAAG

ATAGGAGAGGATATCGTAAGGAAGTGTTGCGGTCTACCTATTGCCATAAA
AACCATGGCATGTACTCTTAGAAATAAAAGAAAGGATGCATGGAAGGATG
CACTTTCGCGCATAGAGCACTATGACATTCACAATGTTGCGCCCAAAGTC
TTTGAAACGAGCTACCACAATCTCCAAGAAGAGGAGACTAAATCCACTTT
5 TTTAATGTGTGGTTTGTTCCTCGAAGACTTCGATATTCCTACTGAGGAGT
TGATGAGGTATGGATGGGGCTTGAAGCTATTTGATAGAGTTTATACGATT
AGAGAAGCAAGAACCAGGCTCAACACCTGCATTGAGCGACTGGTGCAGAC
AAATTTGTTAATTGAAAGTGATGATGTTGGGTGTGTCAAGATGCATGATC
TGGTCCGTGCTTTTGTTCCTGGGTATGTTTCTGAAGTCGAGCATGCTTCT
10 ATTGTCAACCATGGTAATATGCCCCGAGTGGACTGAAAATGATATAACTGA
CTCTTGCAAAAGAATTCATTAACATGCAAGAGTATGTCTAAGTTTCCAG
GAGATTTCAAGTTTCCAAACCTAATGATTTTGAACTTATGCATGGAGAT
AAGTCGCTAAGGTTTCTCAAGACTTTTATGAAGGAATGGAAAAGCTCCA
TGTTATATCATACGATAAAATGAAGTACCCATTGCTTCCTTTGGCACCTC
15 GATGCTCCACCAACATTCGGGTGCTTCATCTCACTAAATGTTTCATTAAG
ATGTTTGATTGCTCTTGTATTGGAAATCTATCGAATCTGGAAGTGCTGAG
CTTTGCTAATTCTCGCATTGAATGGTTACCTTCCACAGTCAGAAATTTAA
AGAAGCTAAGGTTACTTGATCTGAGATTTTGTGATGGTCTCCGTATAGAA
CAGGGTGTCTTGAAAAGTTTAGTCAAACCTTGAAGAATTTTATATTGGAAA
20 TGCATCTGGGTTTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTG
ACAACCTTTCTGCATTAGAATTCGCGTTCCTTAATAACAAGGCTGAAGTG
AAAAATATGTCATTTGAGAATCTTGAACGATTCAAGATCTCAGTGGGACG
CTCTTTTGATGGAAATATCAATATGAGTAGCCACTCATACGAAAACATGT
TGC.AATTGGTGACCAACAAAGGTGATGTATTAGACTCTAACTTAATGGG
25 TTATTTTTGAAAACAAAGGTGCTTTTTTTAAGTGTGCATGGCATGAATGA
TCTTGAAGATGTTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTTCAT
TCTGCAATTTAAAAGTTCCTATTATTTCAAAGTGTGTAGAGTTGAGATAC
CTTTTCAAACCTCAATCTTGCAAACACTTTGTCAAGACTTGAGCATCTAGA
AGTTTGTGAATGCGAGAATATGGAAGAACTCATACATACTGGAATTTGTG
30 GAGAAGAGACAATTACTTTCCCTAAGCTGAAGTTTTTATCTTTGAGTCAA
CTACCGAAGTTATCAAGTTTGTGCCATAATGTCAACATAATTGGGCTACC
ACATCTCGTAGACTTGATACTTAAGGGCATTCCAGGTTTCACAGTCATTT
ATCCGCAGAACAAAGTTGCGAACATCTAGTTTGTGAAGGAAGAGGTAGAT
ATATGTTCTTTATGTTAATACAATTTAAATAATATTTTCAACCAAATTTT
35 CAT.AATATATCTGTAATTTGATTGTATGATGTGTTATTGTTTATATGTGG
CTATTAAGGGATGATTATTTTGCAAGTGTGATTTCCTAAGTTGGAGACAC
TTC.AAATTGATGACATGGAGAACTTAGAAGAAATATGGCCTTGTGAACCTT
AGTGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAAGTGAGTAGCTGTGA
TAAGCTTGTGAATCTATTTCCGCGCAATCCCATGTCTCTGTTGCATCATC
40 TTGAAGAGCTTAAAGTCAAGAATTGCGGTTCCATTGAATCGTTATTCAAC
ATTGACTTGGATTGTGTGCGGTGCAATTGGAGAAGAAGACAACAAGAGCCT
CTTAAGAAGCATCAACATGGAGAATTTAGGGAAGCTAAGAGAGGTGTGGA
GGATAAAAGGTGCAGATAACTCTCATCTCATCAACGGTTTTCAAGCTGTT

GAAAGCATAAAGATTGAAAAATGTAAGAGGTTTAGCAATATATTCACACC
TATCACCGCCAATTTTTATCTGGTGGCACTTTTGGAGATTCAGATAGAAG
GTTGCGGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTTCAATTTAA
CTTTCTTAAGTAATTAAGGACTAACCTCCTGTTTTTTGAATAATAAAGAG
5 GTGGGATGACTAAACTTGGGCATCACAATTGCAACAAAATGTTACAAACC
ATGAAACGTTCAAACCATTTCTTGAATTAAGGTTTCAATACAAGTCATTT
AAAAATATGGCTTAAATTTTTTTATATTTATGTATCAACATGATTTTTCA
TTAGAGATCATTATTATAATAGTAAGTTTAAAGCAATTTAAATTAGAACT
AATTCTAACTTTAGCTAATAAATCGTTATAAATGTAAATAATTACTTTTT
10 AGTGAAATAAGCAACGGATTTAATAAGTTAACAACCTTAAATGTCATTTCC
TAACAAAAAAACTATTTGGTTCAGAAGAACCGTAATTCAAGATAACTAA
AATAAAAAATATTTGACATTCACTAAGAGCATTTTTTTTTCTAAATATGAT
TGCAAATGAATAAACTTAAATTTATACAGAAAAGATTTTTATATATGTT
ATACAAAATTTACAAATTGAAACTGGATATGTTAATTAACGGTTTATAAT
15 TCTGGTATCACAAAGGGATATATAATAAAATATTATTTCTGTAGTCATT
TATAATTGTACTAGTTTATAACCCGTGGGAACCATGAGTTCTAAAATTAG
TTAACTTTTCATAATAAAAATTTATAATTATTATTTATTTTAAATAAATT
ATT.AATTAAGAGATGTATCAAAAATTTAAAGTTATTATAACTTCAAATTT
AACATATAATTAGAAAATATATGATCATAACTTTCCGCAACTCTTCTTTT
20 GTATTAAAATGCCCAGAGAAGCTCTTAGTAYATTTTCTAAATCAAAGTCA
CAAACCTAATGAAGCATATAATTTTGTGAAAATCAATTAGCATTAGGTTT
TAAGAGTCACCAAATTCAAAGAGTAATCCAATGCTTTCATTACCACTATG
GAGAAAATATTTTCTTAGTTTAAATGAAATGAAAACAAACATTCAAACCTA
ATTGTTGCTTACTAAACCAAAGACCCATTACTTAGCCAAGAGTTTAACCA
25 AAAAAAATTACATTCATGTATCATTATTCATGACTAGATATATATGAACA
TGAAGGGAGTTTTTATAGAAAATATAATCATAGATATTCAACATAACTTC
ATGGAATTCCTCAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCA
ACCAAAGAGAAGTTTAGCCTAGCATGGCTAAACTCAAGAAAATAAAATAA
GGATTAGAAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTT
30 GTTCTTGATGTTCTTCTAAGTTCTTCAAGTCTCCAGTTGCTCCTAATAAT
GCAAAGGAGAGCCATTAAATTCGTATGTATTGATCCCTTCAAAGCTGCA
CCAACCTCCCTTAAATAACACTCAAAGCAAAAATGACAAAATTGCCCTG
AAGGACCCTATGCGGGTGCCTTGCGCGGGTGGAGCTGAATACGAAAGGTC
TTTGGTCTTTGTGAGGGTGATGCTGTGCGGGTTAGCTTGTCGCATGCTTC
35 CGCGCGGTTGCGGCACATGTGCACAAGTGATGCATGGTGTGTACGTTCTT
GAGTTTTGAGCCTCCGATGCTTAGTCCATTTGGCCCAATTCGAGTCCAAT
CAGCTTATGACCCATTTTTCTTCAAGTTATCTTCAAGTTATCTTCAAGTT
AAGCCCAAATTGCCTTCTCAAATCATCCATAACTTCACAAAATCGCCCG
TTCATCTTAATCCCGAATGCACAATTATTCTCCTGTCTTCCTTTTAAGCA
40 AGATACCACCTTCTTCATGCTTCATCCATCAATAGTACACTTCATGTATC
ATCTCTACTAGTTATTTAGTCCACAATCCTTATTGTCCTCCAAATTTAAT
TATCTCATTTAGTTCCCGTTCCACTAGTTTCCTTAAAATTTGCAATTAAG
CTCACACAAATATTAAGTACCTGAAATGGTCATAAAATAACAAAAAGGAA

AAT.ATGCATGAAGATTAACATAAATGATGAACGAAATATGCTAAAATAGAC
TATAAAATGAAGTAAATAAAATGAAATTATCGCACTCCGACCACCCTTAT
AGGCTTGTAGTCCATCCACCCTTCATTCTTGTACCAATATGGGATGGAA
ACATCATTAATTAAGCCAAAAAACTAACATATAAGGGGTGAGTGACAAAG
5 GTAAGTACTAAAGATGAAAATAATCCATTTTTYTTGTATATACACAACAC
ACACATAGGGGCAGACGTAGGATTTTCATAGTACAGATTGTTGGTGGCACA
TAAGTGTTGCTGGTGACACTTTTTTTTTCTTTTACGTAGTGGCACAACAG
TAG.AAAAAACGARAAATTCGAAATTTTTTACAATGTGTSTAAAAAAAAYA
GTGGTTGTTGGTGGCCACTATGGACACCAAAGTTGAACTGCCCCCTGCGCGC
10 RCACACACACACACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
ARAGWAWGRRRGAKAKARMCSMSYTTGGGATGTGATACTTCTTTTAGGAA
AATGGAGTTATATCTTTGATATTGTATTTTTTTAATGTAATTTATATATT
TAATCATTTTAGTTTATAAGTTTATTTATTTTGATATGAAAAAAAAGT
CTTTTATACATTGGATTTAACATAAAAAATCCAACAATATTAATCAAAAAG
15 ACC.AMACATGTGGACAMWTATGTATATAAWTAATTCACAATAGTCTTTAG
GAATAGNATTATATATATAATTAATTCTCAATGGTCTTAGGAATAGTAAG
TTCTTATATTTCAAACCTTTNGCCACAATTCTTTGKTACTTWGACACTTY
CCTCTCTCTAATTATATATATATATATATATATATATATATATATACACA
CAC.ACACACACACACTAGATGTGTGCCCGCGCAAAGCAGTGACGTNNNGG
20 AGA.ANACTTTCTTAAGCATAAATAATTATTATATTTTTTTATTGGGTATTA
TAT.AATAAAAAATTACAACCTTTTAAATAAAATATTTATGTTTATACTTTA
TATTTATATTGCTTGTATACTATTAATATAAATAAATTAATATTTATGTCT
AATTTATGAAATGTAAATTAATTTAAATACATGAATTTAATATTTTTAAA
ATTTTCAGTTTGCTTCAAATTGAGTTTCTTAATTATTTTTTTTAATTCAN
25 GTATTCAAACCTTTGGTAAGTATTAAGAATTATTTATGCACAATTGATT
TAT.ACAAAAAACTTTGTAACCTTATACATCTTAAAATTCAAGATATAACTA
ACATGTTTACAATATATATATATATANATATATATATATATATATATAT
ATATATATATATATATAGTAAAGCGCANAGGTCATAGGNANAGANTATTT
TCT.ATTATTCTACGTTTTGCCACAAAAGTTTGAACACTTTGCCACTTTTT
30 GTCCCTCCTTAACCTTTTCAATGTTTTGCGACAAAAGTTCCAAAACCTTTG
CCACTTTGATCATTCTCAACTTTTCACCGCATTAGTTTGTGGAGTTGGC
AGTTTTGGTCCCCCTAACTTCGATATTTTCTCCTGCTAGCCAAAAAGGGT
TCC.AGAGTTTCACANTTTTGGTCCCTGACAATAACCAAATGTGAGATGTC
AAATTTTTGCCACATTAGTTTGTGGAGTTGTCCCTTTTGGTCCCCCACA
35 TTCGATATTCTACTATACGACCTTATTTTTCTCAAATAACAACACGTATA
TTT.AATTACCAATGATAGAAATAGATATCAAATAAAGTATTTGTAACACC
GTGTAAGAACGGTGCTACTATAGGTAAAAATAAACATTTCAAAGTACGAT
GTCCTAATTGGAAAAAGAGTTTTAAAAAAATAACAACCTAGGGGCGAGTTT
TTTTTACAAGTTTGTATCAAATCATATCAAATTTAAGGTGGAACGGTGA
40 CCACATTAACCAGAAATGTAATTTATTCTTTGATTTTGATAATTTTTAAT
ATTTTGTGTGATCTATGTATTTAAAGTAAACAACAAGAACATAATCC
AAAACCCTAAATTGCAAGTCTCGCCCAATTTCTCTATCACTAGTCGTCAC
TTACGATGGCGTTACGTGCTCTCTCACTTCTTACAACCCTTTGTTGCTA

CTCATTACAATAACGAAAAGTTGAATATCCATATATTTATTTGGATGTGG
AATTGAACAAATCTCGTCAAATTTTGGATTTTGTGATGGATTTGAGTAG
AAGTTTGGGCAGAACGGGAATGATGGTCTGCAAGTGGTTATAAACTTGAT
TCTGAGTTATTACTATATATGTAGCCTCTTTACAACGACCAAGGTTTCTT
5 CCAGGTACCATTTGATCTTTTTAGAACCCAGTTGTCTGAAACACCCTGAT
TTGGATCAAATATCACCAACAACCTCTTAAAACTTGATTAATCAATTGTT
TTCTTCATCTTGATAACAAGTGGAATGATTTTCTACTTAGATTAACCTGA
AAAAAAAGGTCCATGTGCGTCTGGTGGATCTGGTAAATGAAGATGGAAGG
GAGAGCTGACTTTAAAGACACAAACACGTCACCATATCTTTTATTTTATT
10 TTAAATTTGCTTTTTTCTTATTCTTTCTTTCTTGATCTCCAGATGGTAT
GTGGTGTGGATAATTTACACATAGAGATTGGGAACGACTGTGTTTTAGAG
AGGACGTGGCTTGGGGTTGAGGATGGTTTATGGCTGGCCGAGTTTCATTT
ATATAAACAAACAAATATATAAAACAAGGGGTAAAATGGCCATCTTATAT
GTATTTAACCGTCCTTTTTTATTTTTTTTTTATTTTTTAAATTTAAGAAGG
15 GGTATACCAAGTGTCAGCCTCTTATTTCCCAACCAGGCAACCAGTCAAATAG
GGACTTAGGTTGTTTGGAAACAGTTCCGTGAGACCGTGACTTGATGGTA
GATAAATTTAGTAACTTAACCCTTCAATTAACCTACCTTTTTCTTATTA
ACTCAATTTCAACCTAAATTCTGATTCTTGTTTGAAAATAAGTTGCATCT
TTATGTTTGTATTATCCTGTTGCATAGGATCCTTAGCATCTTTTAATAAT
20 TTATTTGAAGGTGAAAGATCCAACCTATTTTTTAGCTGTTGGCATTTTCCA
TCATTTGCAACTGTTTCTTGAAAAAAAATACCTAAAATCAAAATAACCA
TTTTCAAATCCAAAATTATAAGAGAGAATTGTTAATGGACGTGGAATCGT
AAATCATTAAACACAGTTCAGTACACAAGTTGCTAATTACATTTCTTGCTG
TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGTCACTGAT
25 ACTAATATTTCTAATGATGTTGTATTATTTCCCATCCTGTCTCATGCACTC
TTTTCATAACCTCCATAAACTTAAATTGGAGAGAGTTAAAGGAGTGGAGG
TGGTGTGTTGAGATAGAGAGTGAGAGTCCAACAAGTAGAGAATTGGTAACA
ACTCACCATAACCAACAACATCCTATTATACTTTCCCAACCTCCAGGAATT
GGATCTAAGTTTTATGGACAACATGAGTCATGTGTGGAAGTGCAGCAACT
30 GGAATAAATTCTTCACTCTTCCAAAACAACATCAGAATCCCCATTCCAC
AACCTCACAACCATAACATGTTTCAGCTGCAGAAGCATTAAAGTACTTGTT
TTCGCTCTCATGGCAGAACTTCTTTCCAACCTAAAGGATATCTGGATAA
GTGGGTGTAATGGTATTAAAGAAGTTGTTTCAAAGAGAGATGATGAGGAT
GAAGAAATGACTACATTTACATCTACCCACACAACCACCATCTTGTTCCC
35 TCATCTTGATTCTCTCACTCTAAGACTACTGGAGAATCTGAAGTGTATTG
GTGGAGGTGGTGCCAAGGATGAGGGGAGCAATGAAATATCTTTCAATAAT
ACC.ACTGCAACTACTGCTGTTCTTGATCAATTTGAGGTATGCTTTGTACA
TATTCAATTATTTATTTAATTTCTTTTTTCTTTGCAATATTCTATAAAT
AATACATTTTATACCCACTATACTAAGATAATAATTACCTAGAGGGATGG
40 ATGCTATGACACAGCTGCTACACTTCAGAACTCTAGTAAGGGCAGTTAT
GGAAGTTCAATAAAATGATAATGGCATCTTTTGATGGGTAATATAGGCAA
TTT.AAGTTTTATTTCTGTAAAGCAGTATTTAGCAAGTACTGGCCAGTAG
GAGAGGAGAATATCACCTTTTGTGAAAATCTGGTCATTGTACCCAAGAAT

TTAGTTAAATGTAACATTTTAGATATCAGGGGACATCAGGTGACAGATAT
TGTAGAATAGAACAATATATAATATTACCCAAAACCTATTTTTTCTAAGGT
TATTCTGTAAATATGTGCTTTCTTGATTTCATTGAATTTGCATTCCCTAT
ATTTTAGGTGGTAAAGTGATTGTCTCTTCAATAAATCCCGAAATTAATTA
5 AAAAAAAAAAAAACAAGTAAATTTTTGATATGGAGAGCACTGGTATCA
TTAGTATATAAAAAAACTAGATTTTGAATTAAGTTTCTTATATAAAAGC
TGTGTATATAGTTTAATTAGTTTTACATCATTTTTCCATGTGGTGTTGCA
GTTGTCTGAAGCAGGTGGTGTCTTCTGGAGTTTATGCCAATACGCTAGAG
AGATAGAGATATCTAAGTGTAATGTATTGTCAAGTGTGATTCCATGTTAT
10 GCAGCAGGACAAATGCAAAGCTTCAAGTGCTGAGAGTAACGGGTGTGA
TGGCATGAAGGAGGTATTTGAAACTCAATTAGGGACGAGCAGCAACAAAA
ACAGAAAGGGTGGTGGTGATGAAGGAAATGGTGGAATCCAAGAGTAAAT
AACAAATGTTATTATGCTTCCCAATCTAAAGACATTGAAAATCTACATGTG
CGGGGGTTTGGAACATATATTCACATTCTCTGCACTTGAAAGCCTGACAC
15 AGCTCCAAGAGTTAAAGATAGTGGGTTGCTACGGAATGAAAGTGATTGTG
AAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACAACAACAACAAC
AACGAAGGGGGCATCTTCTTCTTCTTCTTCTTCTTCTTCTAAGAAGGTTG
TGGTCTTTCCCGTCTAAAGTCCATTGAACTATTCAATCTACCAGAGCTG
GTAGGATTCTTCTTGGGGATGAATGAGTTCCGGTTGCCTTCATTGGAAGA
20 AGTTACCATCAAGTATTGCTCAAAAATGATGGTGTTTGCAGCTGGTGGGT
CCACAGCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACT
CTTGATCAAGAATCTGGCCTTAACCTTTCATCAGGTATATATATATTCCTT
TAATTGGCATGATCTAATTAAGAAAGATATCATTCTGCCAAGTAAATTT
ACTTCAAACACATTACACTGGTTTCAGTCTAAGTTTATGTTGTTCTAGG
25 AAGGCCAAAATGGGAAAGCAAGATAGGGAAAAATAGTGTATTTCAAGTGA
AAGGGTATTTTAGGTATTTTCTGTCAAAAGTTGTTATTGCAGGCTTTTGA
GTACCTGGAATCGTGTGTGGGAGGAGCGTTATTATTCTGATTTGCTTGTT
TCTTTATCATTTTTTCTTAGCCTCTCGAACAGCTAGAAACCCTTTTAATC
TTTTGATTTTAAATGACAAAATTTTTCCCTGTTACTCTATTTGATTGTTG
30 TTCTTCATGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTTCTTTTGAT
TGTTATTTTCATATCATGTTGTCCTTTGAATCAAGCTTTTCCATTTTCAA
CCAGGGCAAAGGTCAAAGTAACCTACTTTATGAGATCAAAAACAGCAA
CCCATCGGATAACTTTTAGTTGGAGTTAATAGTTACAATTACCATTGTGA
TTAATAATTATAATATCTTGTAATTAATTCATTAAAATTGGTACAGCACAT
35 ATATGACATTTTAAAGGTTTGTGTTTTGTTWGACATATATATGCCTCTGGC
GTTTTCTTTATTGGACATGCAGACCTCATTCCAAAGTTTATACGGTGACA
CCTCGGGCCCTGCTACTTCAGAAGGGACAACCTTGGTCTTTTCATAACTTG
ATCGAATTAGATATGGAATTAATTAATGATGTTAAAAAGATTATTCCATC
CAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAGATTATGTGAGTAGTT
40 GTTATTGGGTAGAGGAGGTATTTGAACTGCATTGGAAGCAGCAGGGAGA
AATGGAAATAGTGGAATTGGTTTTGATGAATCGTCACAACTACTACTAC
TACTACTCTTTTCAATCTTCGAAACCTCAGAGAAATGAAGTTGCATTTTC
TACGTGGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTTGAG

TTTCCAAACCTAACAAGAGTTCATATAAGTAGGTGTAGAAGGTTAGAACA
TGTATTTACTAGTTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAG
ATATTAGTTGGTGCAACCATATGGAGGAGGTGATTGTAAAGGATGCAGAT
GTTTCTGTTGAAGAAGACAAAGAGAGAGAATCTGATGGCAAGACGAATAA
5 GGAGATACTTGTGTTACCTCGTCTAAAATCCTTGAAATTAAAATGCCTTC
CATGTCTTAAGGGGTTTAGCTTGGGGAAGGAGGATTTTTCATTCCCATT
TTGGATACTTTAGAAATCTACAAATGCCAGCAATAACGACCTTCACCAA
GGGAAATTCTGCTACTCCACAGCTAAAAGAAATAGAAACAAGATTGGCT
CGTTTTATGCAGGGGAAGACATCAACTCCTCTATTATAAAAAGATCAAAC
10 AACAGGTAAATCAGATCTTTGTTGCTTAATAATTCTTAAACTACATTG
AAAAGCTTCATGCAAGTTTTTTTTTGTATATTGTCAAAAACCGCAACCTA
CATTTTCAGCTTTATATTTATGTACTTTATGCAGGAGTTCAAACAAAAC
CTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTTCATGTTCT
AGTTGCCTATTAATTAATGGCCTTTTAGTTCRTGATTTTTGGATGTAGTY
15 WTCATGATGATGTGAATCTTCTAATACCCCATTCATTGTTTGGTTGAATG
TTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTTCATCATATG
AAGGACATTAAAGAACATGGATGCTATGAAGATGTTGGAARAC

RG2S deduced polypeptide sequence (SEQ ID NO:125)

20 MSDPTGIAGAIINPIAQRALVPVTDHVGYMISCRKYVRVMQTKMTELNTSRISVEEH
ISRNTRNHLQIPSQIKDWLDQVEGIRANVENFPIDVITCCSLRIRHKLKGQKAFKITEQI
ESLTRQLSLISWTDDPVPLGRVGS MNASTSASSSDDFPSREKTFTQALKALEPNQQF
HMYALCGMGGVGKTRMMQRLKKAEEKKLFNYIVRAVIGEKTDPFAIQEAIADYL
GIQLNEKTKPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGLSPFPNQG
25 VDFKVLLTSRDSQVCTMMGVEANSIINVGLLTEAEAQSLFQQFVETSEPELQKIGED
IVRKCCGLPIAIKTMACTLRNKRKDAWKDALSRIEHYDIHNVAPKVFETSYHNLQE
EETKSTFLMCGLFPEDFDIPEELMRYGWGLKLFDRVYTIREARLNTCIERLVQT
NLLIESDDVGCVKMHDLVRAFVLGMFSEVEHASIVNHGNMPEWTENDITDSCKRIS
LTCKSMSKFPDGFKFPNLMILKLMHGDKSLRFPQDFYEGMEKLHVISYDKMKYPLL
30 PLAPRCSTNIRVLHLTKCSLKMFD CSCIGNLSNLEVLSFANSRIEWLPSTVRNLKKLR
LLDLRFCDGLRIEQGVLSLVKLEEFYIGNASGFIDDNCNEMAERSDNLSALEFAFF
NNKAEVKNMSFENLERFKISVGRSFDGNINMSSHSYENMLQLVTNKGDVLD SKLN
GLFLKTKVLFLSVHGMNDLEDVEVKSTHPTQSSSFCNLKVLISKVELRYLFLKLN
ANTLSRLEHLEVCECENMEELIHTGICGEETITFPKLKFLSLSQLPKLSSLCHNVNIIG
35 LPHLVDLILKGIPGFTVIYPQNKLR TSSLLKEEVVIPKLETQIDD MENLEEIWPCELS
GGEKVKLREIKVSSCDKLVNLFPRNPMSLLHHLEELKVKNCGSIESLFNIDLDCVGA
IGEEDNKSLRSINMENLGKLREVWRIKGADNSHLINGFQAVESIKIEKCKRFSNIFT
PITANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEVTD TNISNDVVLFP SCLMH
SFHNLHKLKLERVKGVEVVFEIESESPTSRELVTTHNQQHPIILPNLQELDLSFMD
40 NMSHVWKC SNWNKFFTL PKQQSES PFHNLTTIHMFS CRSIKYLF SPLMAELLSNLK
DIWISGCNGIKEVVSKR DDEDEEMTTFTSTHTTTILFPHLDSLTLRLLENLKCIGGGG
AKDEGSNEISFNNTTATTAVLDQFELSEAGGVSWSLCQYAREIEISKCNVLSSVIPCY

AAGQMQLQVLRVTGCDGMKEVFETQLGTSSNKNRKGGGDEGNNGGIPRVNNNVI
MLPNLKTLLKIYMCGGLEHIFTFSALESLTQLQELKIVGCYGMKVIVKKEEDEYGEQ
QTTTTTTTKGASSSSSSSSSKKVVFPRLKSIELFNLPELVGFFLGMNEFRLPSLEEV
IKYCSKMMVFAAGGSTAPQLKYIHTRLGKHTLDQESGLNFHQTSFQSLYGDTSGPA
5 TSEGTTWSFHNLIELDMELNYDVKKIIPSELLQLQKLEKIHVSSCYWVEEVFETAL
EAAGRNGNSGIGFDESSQTTTTTTLFNLRLNREMKLHFLRGLRYTWKSNQWTAFEF
PNLTRVHISRCRRLEHVFTSSMVGSLLQLQELDISWCNHMEEVIVKDADVSVEEDK
ERESDGKTNKEILVLPRLKSLKLKCLPCLKGFSLGKEDFSFPLDLEIYKCPAITTFT
10 KGN SATPQLKEIETRFGSFYAGEDINSSIIKRSNNRSSNKT LINVK.ILK

RG2T polynucleotide sequence (SEQ ID NO:126)

GGAAGACGACAATGGTGCAACGGTTGAAGAAGGTTGTGAAAGATAAGAAG
ATGTTCCATTATATTGTCGAGGTGGTTGTAGGGGCAAACACTGACCCCAT
TGCTATCCAGGATACTGTTGCAGATTACCTCAGCATAGAACTGAAAGGAA
15 ATACGAGAGATGCAAGGGCTTATAAGCTTCGTGAATGCTTTAAGGCCCTC
TCTGGTGGAGGTAAGATGAAGTTCCTAGTAATTCTTGACGATGTATGGAG
CCCTGTTGATCTGGATGATATCGGTTTAAGTTCTTTGCCAAATCAAGGTG
TTG.ACTTCAAGGTCTTGCTGACATCACGCAACAGTGATATCTGCATGATG
ATGGGAGCTAGTTTAATTTTCAACCTCAATATGTTAACAGACGAGGAAGC
20 ACATAATTTTTTCCGTCGATACGCAGAAATTTCTTATGATGCTGATCCCG
AGCTTATTAAGATAGGAGAAGCTATTGTAGAGAAATGTGGTGGTTTACCC
ATTGCCATCAAACTATGGCCGTTACTCTTAGAAATAAACGCAAAGATGC
ATGGAAAGATGCACTTTCTCGTTTAGAGCACCGTGACACTCATAATGTTG
TGGCTGATGTTCTTAAATTGAGCTACAGCAATATCCAAGACGAGGAGACT
25 CGGTCGATTTTTTTGCTATGTGGTTTGTTCCTGAAGACTTTGATATTCC
TACCGAAGACTTAGTGAGGTATGGATGGGGATTGAAAATATTTACCAGAG
TGT.ATACTATGAGACATGCAAGAAAAAGGTTGGACACGTGCATTGAGCGG
CTT.ATGCATGCCAACATGTTGATAAAAAGTGATAATGTTGGATTTGTCAA
GATGCATGATCTGGTTCGTGCTTTTGTGTTTGGGCATGTTATCTGAAGTCG
30 AGCATGCATCAATTGTCAACCATGGGGATATGCCAGGGTGGTTTGAAACT
GCAAATGATAAGAACAGCTTGTGCAAAAAGAATTTCATTAACATGCAAAGG
TATGTCTGCGATTCTCTGAAGACCTCACGTTTCCAAACCTCTCGATCCTGA
AATTAATGGATGGAGACGAGTCACTGAGGTTTCTGAAGGCTTTTATGGA
GAAATGGAAAACCTTCAGGTTATATCATATGATAACATGAAGCAGCCATT
35 TCTTCCACAATCACTTCAATGCTCCAATGTTTCGAGTGCTTCATCTCCATC
ACTGCTCATTAAATGTTTGATTGCTCTTCTATTGGAAATCTTTTGAATCTC
GAGGTGCTCAGCATTGCTAATTCTGCCATTAAATTGTTACCCTCCACTAT
TGGAGATCTGAAGAAGCTAAGGCTCCTGGATTTGACAAATTGTGTTGGTC
TCTGTATAGCTAATGGCGTCTTTAGAAATTTGGTCAAACCTTGAAGAGCTT
40 TAT.ATGAGAGTTGATGATCGAGATTCGTTTTTTGTGAAAGCTGATGACAG
CAAGACCATTACCT

RG2T deduced polypeptide sequence (SEQ ID NO:127)

KTMMVQRLKKVVKDKKMFHYIVEVVVGANTDPIAIQDTVADYLSIELKGNTRDAR
AYKLRECFKALSGGGKMKFLVILDDVWSPVDLDDIGLSSLPNQGVDFKVLLTSRNS
DICMMMGASLIFNLNMLTDEEAHNFFRRYAEISYDADPELIKIGEAIWEKCGGLPIAI
5 KTMAVTLRNKRKDAWKDALSRLEHRDTHNVVADV LKLSYSNIQDEETRSIFLLCG
LFPEDFDIPTEDLVRYGWGLKIFTRVYTMRHARKRLDTCIERLMHANMLIKSDNVG
FVKMHDLVRAFVLGMLSEVEHASIVNHGDMPGWFETANDKNSLCKRISLTCKGMS
AIPEDLTFPNLSILKLMDGDESLRFPEGFYGEMENLQVISYDNMKQPFLPQSLQCSN
VRVLHLHHCSLMFDCSSIGNLLNLEVLSIANSIAIKLLPSTIGDLKKLRLDLTNCVGL
10 CIANGVFRNLVKLEELYMRVDDRDSFFVKADDSKTIT

RG2U polynucleotide sequence (SEQ ID NO:128)

GCCTTGTGTGGGATGGGTGGAGTGGGAAAGACCACTGTGATGAAGAAGCT
GAAGGAGGTTGTGGTAGGAAAGAAACTGTTTAATCATTATGTTGAGGCGG
15 TTATAGGGGAAAAGACAGACCCCATTTGCTATTCAACAAGCTGTTGCCGAG
TACCTTGGTATAAGTCTAACCGAAACCACTAAACCAGCAAGAAGTGTAA
GCTCCGTACATGGTTTGCAAACAACCTCAAATGGAGGAAAGAAGAAGTTCC
TGGTAATACTAGACGATGTATGGCAACCAGTTGATTTGGAAGATATTGGT
TTAAGTCGTTTTCCAAATCAAGATGTTGACTTCAAGGTCTTGATTACATC
20 ACGGGACCAATCAGTTTGCACTGAGATGGGAGTTAAAGCTGATTTAGTTC
TCAAGGTGAGTGTCTTGGAGGAAGCGGAAGCACACAGTTTGTTCTCTCAA
TTTTTAGAACCTTCTGATGATGTCGATCCTGAGCTCAATAAAATCGGAGA
AGAAATTGTAAAGAAGTGTTCGAGACTACCCATTGCTATCAAAACCATGG
CCTGAACCTTAGAAGTAAAAGTAAGGATACATGGAAGAATGCCCTTTCT
25 CGTTTACAACACCATGACATTAACACAATTGCGTCTACTGTTTTCCAAAC
TAGCTATGACAATCTCGAAGACGAGGTGACTAAAGCTACTTTTTTGCTTT
GTGGTTTATTTCCGGAGGACTTCAATATTCCTACCGAGGACCTATTGAGG
TATGGATGGGGATTGAAGTTATTCAAGGAAGTAGATACTATACGAGAAGC
AAGATCCAAGTTGAAAGCCTGCATTGAGCGGCTCATGCATACCAATTTGT
30 TGATCGAAGGTGATGATGTTAGGTACGTAAAGATGCATGATCTGGTGCGT
GCTTTTGTGTTTGGATATGTTTTCTAAAGCCGAGCATGCATCTATTGTCAA
CCATGGTAGTAGTAAGCCAAGGTGGCCTGAAACTGAAAGTGATGTGAGCT
CCTCTTGCAAAGAATTTCAATTAACATGCAAGGGTNTG

35 RG2U deduced polypeptide sequence (SEQ ID NO:129)

ALCGMGGVGKTTVMKKLKEVVVGKKLFNHYVEAVIGEKTDPPIAIQQA VAEYLGIS
LTETTKPARTDKLRTW FANN SNGGKKKFLVILDDVWQPVDLEDIGLSRFPNQDVD
FKVLITSRDQSVCTEMGVKADLV LKVS VLEEAEAHSLFLQFLEPSDDVDPELNKIGE
EIVK KCCRLPIAIKTMA.TLRSKSKDTWKNALSRLQHHDINTIASTVFQTSYDNLEDE
40 VTKATFLLCGLFPEDFNIPTEDLLRYGWGLKLFKEVD TIREARSKLKACIERLMHTN

LLIEGDDVRYVKMHDLVRAFVLDMFSKAEHASIVNHGSSKPRWPETESDVSSSCKR.
ISLTCKG?

RG2V polynucleotide sequence (SEQ ID NO:130)

5 CTGTGGAAGACACGAATGATSAAGAAGCTGAAGGAGGTCGTGGAACAAAA
GAAAATGTTCAATATTATTGTTCAAGTGGTCATAGGAGAGAAGACAAACC
CTATTGCTATTCAGCAAGCTGTAGCAGATTACCTCTCTATTGAGCTGAAA
GAAAACACTAAAGAAGCAAGAGCTGATAAGCTTCGTNAATGGTTCGAGGA
CGATGGAGGAAAGAATAAGTTCCTTGTAATACTTGATGATGTATGGCAGT
10 TTGTCGATCTTGAAGATATTGGTTTAAAGTCCTCTGCCAAATAAAGGTGTC
AACTTCAAGGTCTTGTTGACGTTAAGAGATTCACATGTTTGCCTCTGAT
GGGAGCTGAAGCCAATTCAATTCTCAATATAAAAGTTTTAAAGATGTTN
AAGGACAAAGTTTGTTCGCCAGTTTGCTAAAAATGCAGGTGATGATGAC
CTGGATCCTGCTTTCAATGGGATAGCAGATAGTATTGCAAGTAGATGTCA
15 AGGTTTGCCCATTGCCATCAAAACCATTGCCTTAAGTCTTAAAGGTAGAA
GCAAGCCTGCGTGGGACCATGCGCTTCTCGTTTGGAGAACCATAAGATT
GGTAGTGAAGAAGTTGTGCGTGAAGTTTTTAAAATTAGCTATGACAATCT
CCAAGATGAGGTTACTAAATCTATTTTTWTACTTTGTGCTTTATTTCTG
AAGATTTTGATATTCCTATTGAGGAGTTGGTGAGGTATGGGTGGGGCTTG
20 AAATTATTTATAGAAGCAAAAACCTATAAGAGAAGCAAGAAACAGGCTCAA
CACCTGCACTGAGCGGCTTAGGGAGACAAATTTGTTATTTGGAAGTGATG
ACATTGGATGCGTCAAGATGCACGATGTGGTGCGTGATTTTGTGTTGGTAT
ATATTCTCAGAAGTCCAGCACGCTTCAATTGTCAACCATGGTAATGTGTC
AGAGTGGCTAGAGGAAAATCATAGCATCTACTCTTGTAAGAATTTTCAT
25 TAACATGCAAGGGTATGTCTGAGTTTCCCAAAGACCTCAAATTTCCAAAC
CTTTCAATTTTGAACTTATGCATGGAGATAAGTCGNTGAGCTTTCTCTGA
AGACTTTTATGGAAAGATGGAAAAGGTTCAAGGTAATATCATATGATAAAT
TGATGTATCCATTGCTTCCCTCATCACTTGAATGCTCCACTAACGTTCTGA
GTGCTTCATCTCCATTATTGTTTATTAAAGGATGTTTGATTGCTCTTCAAT
30 TGGTAATCTTCTCAACATGGAAGTGCTCAGCTTTGCTAATTCTAACATTG
AATGGTTACCATCTACAATTGGAAATTTGAAGAAGCTAAGGCTACTAGAT
TTGACAAATTGTAAAGGTCTTCGTATAGATAATGGTGTCTTAAAAAATTT
GGTCAAACCTGAAGAGCTTTATATGGGTGTTAATGTCCGTATGGACCAGG
CCGT

35

RG2V deduced polypeptide sequence (SEQ ID NO:131)

LWKTRM?KKLKEVVEQKKMFNIIQVQVIGEKTNPQIAIQQAVADYLSIELKENTKEAR
ADKLR?WFEDDGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTLRDSH
VCTLMGAEANSILNIKVLKDV?GQSLFRQFAKNAGDDDLDPAFNGIADSIASRCQGL
40 PIAIKTIALSLKGRSKPAWDHALSRLNHNKIGSEEVVREVFKISYDNLQDEVTKSIF?L
CALFPEDFDIPIEELVRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIG

CVKMHDVVRDFVWYIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEF
PKDLKFPNLSILKLMHGDKS?SFPEDFYGKMEKVQVISYDKLMYPLLPSSECTNV
RVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKLRLLDLTNCKG
LRIDNGVLKNLVKLEELYMGVNVVRMDQAV

5

RG2W polynucleotide sequence (SEQ ID NO:132)

TTGGGAAAGAGACAATGATGAAGAATTGAAAGAGGTTGTGGTTGAAAAGA
AAATGTTTAATCATTATGTGGAGGCGGTTATAGGGGAGAAGACGGACCCC
ATTGCTATTCAGCAAGCCGTTGCAGAGTACCTTGGTATAATTCTAACAGA
10 AACCATAAGGCAGCAAGAACCGATAAGCTACGTGCATGGCTTTCTGACA
ATTCAGATGGAGGAAGAAAGAAGTTCCTAGTAATACTAGACGATGTATGG
CATCCGGTTGATATGGAAGATATTGGTTTAAGTCGTTTCCCAAATCAAGG
TGTCGACTTCAAGGTCTTGATTACATCACGGGACCAAGCTGTTTGCCTG
AGATGGGAGTTAAAGCTGATTCAGTTATCAAGGTGAGTGTCTAGAGGAA
15 GCTGAAGCACAAAGCTTATTCTGCCAACTTTGGGAACCTTCTGATGATGT
CGATCCTGAGCTCCATCAGATTGGAGAAGAAATTGTAAGGAAGTGTGTG
GTTTACCCATTGCAATAAAAACCATGGCCTGCACTCTTAGAAGTAAAAGC
AAGGATACATGGAAGAATGCACTTTCTCGTTTACAACACCATGACATTAA
CACAGTCGCGCCTACTGTTTTTCAAACAGCTATGACAATCTCCAAGATG
20 AGGTGACTGGAGATACTTTTTTGCTATGTGGTTTGTTCGGAGGACTTC
GATATTCCTACTGAAGACTTATTGAAGTATGGATGGGGCTTAAATTATT
CAAGGGAGTGGATTCTGTAAAGAGAAGCAAGATAACAGTTGAACGCCTGCA
TTGAGCGGCTCGTGCATACCAATTTGTTGATTGAAAGTGATGTTGTTGGG
TGCGTCAAGTTGCACGATCTGGTGCGTGCTTTTATTTTGGATATGTTTTG
25 TAAAGCGGAGCATGCTTCGATTGTCAACCATGGTAGTAGTAAGCCTGGGT
GGCCTGAAACTGAAAATGATGTGATCAGGACCTCCTGCAAAAGAATCTCA
TTAACATGCAAGGGTATGATTGAGTTTTCTAGTGACCTCAAGTTTCCAAA
TGTCTTGATTTTAAACTTATGCATGGAGATAAGTCGCTAAGGTTT

30 RG2W deduced polypeptide sequence (SEQ ID NO:133)

WERDNDEELKEVVVEKKMFNHVEAVIGEKTDPPIAIQQA VAEYLGII LTETTKAAR
TDKLRAWLSDNSDGGRRKKFLVILDDVWHPVDMEDIGLSRFPNQGVDFKVLITSRD
QAVCTEMGVKADSVIKVSVLEEA EAQSLFCQLWEP SDDVDPELHQIGEEIVRKCCG
LPIAIKTMACTLRSSKSDTWKNALSRLQHHDINTVAPT V FQTSYDNLQDEVTGDTF
35 LLCGLFPEDFDIPTEDLLKYGWGLKLFKGVDSVREARYQLNACIERLVHTNLLIESD
VVGCVKLHDLVRAFILDMFCKAEHASIVNHGSSKPGWPETENDVIRT SCKRISLTCK
GMIEFSSDLKFPNVLILKLMHGDKSLRF

RG5 polynucleotide sequence (SEQ ID NO:134)

40 GGGGGGGTGGGGAAGNCGACTCTAGCCCAGAAGNTCTATAATGACCATAA
AATAAAAGGAAGCTTTAGTAAACAAGCATGGATCTGTGTTTCTCAACAAT

ATTCTGATATTTTCAGTTTTTGAAAGAAGTCCTTCGGAACATCGGTGTTGAT
TATAAGCATGATGAAACTGTTGGAGAACTTAGCAGAAGGCTTGCAATAGC
TGTCGAAAATGCAAGTTTCTTTCTTGTGTTGGATGATATTTGGCAACATG
AGGTGTGGACTAATTTACTCAGAGCCCCATTAAACACTGCAGCTACAGGA
5 ATAATTCTAGTAACAACCTCGTAATGATACAGTTGCACGAGCAATTGGGGT
GGAAGATATTCATCGAGTAGAATTGATGTCAGATGAAGTAGGATGGAAAT
TGCTTTTGAAGAGTATGAACATTAGCAAAGAAAGTGAAGTAGAAAACCTA
CGAGTTTTAGGGGTGACATTGTTTCGTTTGTGTGGTGGCCTCCCCCTAGC
CTT

10 **RG5 deduced polypeptide sequence (SEQ ID NO:135)**

GGVGKTTLAQK?YNDHKIKGSFSKQAWICVSQQYSDISVLKEVLRNIGVDYKHDET
VGELSRRLAIAVENASFFLVLDIHWQHEVWTNLLRAPLNTAATGILVTTRNDTVA
RAIGVEDIHRVELMSDEVGWKLLLKSMNISKESSEVENLRVLGVDIVRLCGGLPLAL

15 **RG7 polynucleotide sequence (SEQ ID NO:136)**

GGTGGGGTTGGGAAGACAACGGGCACAAGGAGGCGACTGCCAATACTTCC
GACTTTTATTCATAGAGATGACGAGTCTTATTTTCCTACTACTATAGGGA
GGATATTTGGTTGCGCGAGACGATTTCATTGCGCGAAGGGATTCTATCCTT
20 CTTTTTTTCCGCGAAGACTTCGTTCCGGAGGACGGGCTATATTCCCTTTA
ATATTAGTCTAGCCCAGTCTAGGCCAACCATATGGCGATGCGGTAGACCT
CCCAGAGATAGATACTTGATCTTAGAGGATTCACACGTTCAATGGTGGAA
ACTTAAGGAACCGGCTAAGAGTGACTAAACGGAAAAACCCTATTCATTCC
ATAGCCTCATCCGGTCGAGGCATTAAACAATCCATCCCAATCCTCTTTCC
25 TTTGGTCTACTCTAATGATGTGCCCCGTTTCGTTGGTGAATATCTCTTTAT
ACCGACGATTTATATGGGGATTGCCACTAGCGTTG

30 The above examples are provided to illustrate the invention but not to limit
its scope. Other variants of the invention will be readily apparent to one of ordinary skill
in the art and are encompassed by the appended claims. All publications, patents, and
patent applications cited herein are hereby incorporated by reference.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid construct comprising an RG polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide from an RG family selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

2. The nucleic acid construct of claim 1, wherein the RG polynucleotide encodes an RG polypeptide comprising an leucine rich region (LRR).

3. The nucleic acid construct of claim 1, wherein the RG polynucleotide encodes an RG polypeptide comprising a nucleotide binding site (NBS).

4. The nucleic acid construct of claim 1, wherein the polynucleotide is a full length gene.

5. The nucleic acid construct of claim 1, wherein the further encodes a fusion protein.

6. The nucleic acid construct of claim 1, wherein the RG1 polypeptide is encoded by an RG1 polynucleotide sequence.

7. The nucleic acid construct of claim 6, wherein the RG1 polypeptide is encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).

8. The nucleic acid construct of claim 1, wherein the RG2 polypeptide is encoded by an RG2 polynucleotide sequence.

9. The nucleic acid construct of claim 8, wherein the RG2 polypeptide is encoded by a polynucleotide sequence selected from the group consisting of: SEQ ID NO:21 (RG2A);

SEQ ID NO:23 (RG2B); SEQ ID NO:25 (RG2C); SEQ ID NO:27 (RG2D); SEQ ID
NO:29 (RG2E); SEQ ID NO:31 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:35
(RG2H); SEQ ID NO:37 (RG2I); SEQ ID NO:39 (RG2J); SEQ ID NO:41 (RG2K); SEQ
ID NO:43 (RG2L); SEQ ID NO:45 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89
5 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D);
SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID
NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107
(RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L);
SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID
10 NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126
(RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132
(RG2W).

10. The nucleic acid construct of claim 1, wherein the RG3 polypeptide is encoded by
15 an RG3 polynucleotide sequence.

11. The nucleic acid construct of claim 10, wherein the RG3 polypeptide is encoded by
a polynucleotide sequence as set forth in SEQ ID NO:68.

20 12. The nucleic acid construct of claim 1, wherein the RG4 polypeptide is encoded by
an RG4 polynucleotide sequence.

13. The nucleic acid construct of claim 12, wherein the RG4 polypeptide is encoded by
a polynucleotide sequence as set forth in SEQ ID NO:69.

25

14. The nucleic acid construct of claim 1, wherein the RG5 polypeptide is encoded by
an RG5 polynucleotide sequence.

15. The nucleic acid construct of claim 14, wherein the RG5 polypeptide is encoded by
30 a polynucleotide sequence as set forth in SEQ ID NO:134.

16. The nucleic acid construct of claim 1, wherein the RG7 polypeptide is encoded by an RG7 polynucleotide sequence.

17. The nucleic acid construct of claim 16, wherein the RG7 polypeptide is encoded by a polynucleotide sequence as set forth in SEQ ID NO:136.

18. The nucleic acid construct of claim 1, further comprising a promoter operably linked to the RG polynucleotide.

19. The nucleic acid construct of claim 18, wherein the promoter is a plant promoter.

20. The nucleic acid construct of claim 19, wherein the plant promoter is a disease resistance promoter.

21. The nucleic acid construct of claim 19, wherein the plant promoter is a lettuce promoter.

22. The nucleic acid construct of claim 18, wherein the promoter is a constitutive promoter.

23. The nucleic acid construct of claim 18, wherein the promoter is an inducible promoter.

24. The nucleic acid construct of claim 18, wherein the promoter is a tissue-specific promoter.

25. A nucleic acid construct comprising a promoter sequence from an RG gene linked to a heterologous polynucleotide.

26. A transgenic plant comprising a recombinant expression cassette comprising a promoter operably linked to an RG polynucleotide.

27. The transgenic plant of claim 26, wherein the plant promoter is a plant promoter.
28. The transgenic plant of claim 26, wherein the plant promoter is a viral promoter.
- 5 29. The transgenic plant of claim 26, wherein the plant promoter is a heterologous promoter.
30. The transgenic plant of claim 26, wherein the plant is lettuce.
- 10 31. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).
- 15 32. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:21 (RG2A); SEQ ID NO:23 (RG2B); SEQ ID NO:25 (RG2C); SEQ ID NO:27 (RG2D); SEQ ID NO:29 (RG2E); SEQ ID NO:31 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:35 (RG2H); SEQ ID NO:37 (RG2I); SEQ ID NO:39 (RG2J); SEQ ID NO:41 (RG2K); SEQ ID NO:43 (RG2L); SEQ ID NO:45 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).
- 20 33. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:68 (RG3) and SEQ ID NO:69 (RG4).
- 25 30

34. The transgenic plant of claim 26, wherein the RG polynucleotide comprises a sequence as set forth in SEQ ID NO:134 (RG5).

35. The transgenic plant of claim 26, wherein the RG polynucleotide comprises a sequence as set forth in SEQ ID NO:136 (RG7).

36. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO: 13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), and SEQ ID NO:20 (RG1J).

37. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

38. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG3 polypeptide with a sequence as set forth by SEQ ID NO:138.

39. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG4 polypeptide with a sequence as set forth by SEQ ID NO:139.

40. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG5 polypeptide with a sequence as set forth by SEQ ID NO:135.

41. A method of enhancing disease resistance in a plant, the method comprising
5 introducing into the plant a recombinant expression cassette comprising a promoter functional in the plant and operably linked to an RG polynucleotide sequence.

42. The method of claim 41, wherein the plant is a lettuce plant.

10 43. The method of claim 41, wherein the RG polynucleotide encodes an RG polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51
15 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O);
20 SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

44. The method of claim 41, wherein the RG polynucleotide encodes an RG polypeptide
25 selected from the group consisting of SEQ ID NO:138 (RG3); SEQ ID NO:139 (RG4); and SEQ ID NO:135 (RG5).

45. The method of claim 41, wherein the promoter is a tissue-specific promoter or a plant disease resistance promoter.

46. The method of claim 41, wherein the promoter is a constitutive promoter or an inducible promoter.

5 47. A method of detecting RG resistance genes in a nucleic acid sample, the method comprising:

contacting the nucleic acid sample with an RG polynucleotide to form a hybridization complex; and,

wherein the formation of the hybridization complex is used to detect the RG resistance gene in the nucleic acid sample.

10 48. The method of claim 47, wherein the RG polynucleotide is an RG1 polynucleotide.

49. The method of claim 47, wherein the RG polynucleotide is an RG2 polynucleotide.

15 50. The method of claim 47, wherein the RG polynucleotide is an RG3 polynucleotide, an RG4 polynucleotide, an RG5 polynucleotide or an RG7 polynucleotide.

51. The method of claim 47, wherein the RG resistance gene is amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide.

20 52. The method of claim 51, where the RG resistance gene is amplified by the polymerase chain reaction.

53. The method of claim 47, wherein the RG polynucleotide is labeled.

25 54. An RG polypeptide having at least 60% sequence identity to a polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US98/00615

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : 435/6, 91.2, 418, 419; 530/350; 536/23.1, 23.6, 24.1; 800/205

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.2, 418, 419; 530/350; 536/23.1, 23.6, 24.1; 800/205

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARAN et al. Development of Reliable PCR-Based Markers Linked to Downy Mildew Resistance Genes in Lettuce. Theor. Appl. Genet. 1993. Vol. 85, No. 8, pages 985-993, see entire article.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	KESSELI et al. Analysis of a Detailed Genetic Linkage Map of Lactuca sativa (Lettuce) Constructed From RFLP and RAPD Markers. Genetics. April 1994. Vol. 136, No. 4, pages 1435-1446, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	MICHELMORE, RW. Isolation of Disease Resistance Genes from Crop Plants. Current Opinion in Biotechnology. 1995. Vol. 6, No. 2, pages 145-152, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54



Further documents are listed in the continuation of Box C.



See patent family annex.

Special categories of cited documents:	
A document defining the general state of the art which is not considered to be of particular relevance	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
B earlier document published on or after the international filing date	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
O document referring to an oral disclosure, use, exhibition or other means	*A* document member of the same patent family
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

14 MARCH 1998

Date of mailing of the international search report

13 APR 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized official

PHUONG BUI

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim N .
Y	PARAN et al. Recent Amplification of Triose Phosphate Isomerase Related Sequences in Lettuce. Genome. 1992. Vol. 35, No. 4, pages 627-635, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	PARAN et al. Identification of Restriction Fragment Length Polymorphism and Random Amplified Polymorphic DNA markers linked to Downy Mildew Resistance Genes in Lettuce, Using Near-Isogenic Lines. Genome. 1991. Vol. 34, No. 6, pages 1021-1027, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US98/00615

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 7, 9, 11, 13, 15, 17, 31-40, 43-44
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

these claims are drawn to numerous sequences identified by SEQ ID NOs. However, since no computer readable form was submitted, no meaningful search could be carried out.
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐

The additional search fees were accompanied by the applicant's protest.

☐

No protest accompanied the payment of additional search fees.

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

A01H 1/00; C07H 21/04; C07K 14/00; C12N 5/04, 5/10; C12P 19/34; C12Q 1/68